

AD-A265-099



■ AGARD-CP-528

1

AGARD-CP-528

# AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

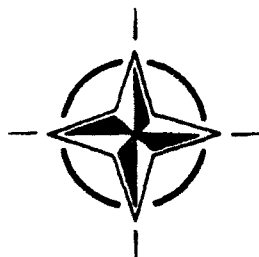
AGARD CONFERENCE PROCEEDINGS 528

## Radiolocation Techniques

(Les Techniques de Radiolocalisation)

DTIC  
ELECTE  
MAY 26 1993  
S C D

*Papers presented at the Electromagnetic Wave Propagation Panel Symposium,  
held in London, United Kingdom, 1st-5th June 1992.*



NORTH ATLANTIC TREATY ORGANIZATION

DISTRIBUTION STATEMENT X

Approved for public release  
Distribution Unlimited

Published November 1992

Distribution and Availability on Back Cover

# AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

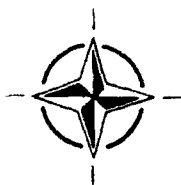
## AGARD CONFERENCE PROCEEDINGS 528

### Radiolocation Techniques

(Les Techniques de Radiolocalisation)

Accession For	
NTIS CRAFT	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and/or Special
A-1	

Papers presented at the Electromagnetic Wave Propagation Panel Symposium, held in London, United Kingdom, 1st—5th June 1992.



North Atlantic Treaty Organization  
Organisation du Traité de l'Atlantique Nord

93-11709



93 5 25 14 6

# The Mission of AGARD

According to its Charter, the mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

- Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community;
- Providing scientific and technical advice and assistance to the Military Committee in the field of aerospace research and development (with particular regard to its military application);
- Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;
- Improving the co-operation among member nations in aerospace research and development;
- Exchange of scientific and technical information;
- Providing assistance to member nations for the purpose of increasing their scientific and technical potential;
- Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Programme and the Aerospace Applications Studies Programme. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.

The content of this publication has been reproduced  
directly from material supplied by AGARD or the authors.

Published November 1992

Copyright © AGARD 1992  
All Rights Reserved

ISBN 92-835-0695-2



*Printed by Specialised Printing Services Limited  
40 Chigwell Lane, Loughton, Essex IG10 3TZ*

# Theme

Modern defence systems must have a comprehensive surveillance capability. A key component of a surveillance system is the ability to detect and to locate radio signals. Requirements are both for global surveillance and for local areas in situations of rapid deployment of forces. The entire radio spectrum is of interest: long wave radio waves are used for strategic communications; the HF band remains one of the important frequency regions used by large and small countries for global and regional communications; the frequency range above 50 MHz through to the millimetre wave region is extensively used for both communications and surveillance purposes. Effects of the propagation medium play an important role in applying location techniques. For HF frequencies and below, the primary medium is the ionosphere; for higher frequencies important environmental effects are due to atmospheric refraction. In the entire radio band both natural and man-made noise (including jamming) often play an important role in direction finding and location applications.

Radiolocation techniques are concerned with the angle of arrival of a signal and location of an emitter. This can be done from a single site (fixed or mobile, including satellites) or a network of stations. Some of the techniques which have been successfully used are angle of arrival, time difference of arrival and doppler direction of arrival measurements. The next generation radiolocation systems will have to use smaller antenna arrays, be more mobile, provide more accurate solutions in real time, be able to fuse data, from various systems, automatically account for the propagation environment in a manner transparent to the user, and will use expert systems to interpret the data and communicate with the user.

## Topics Covered

- General aspects: Propagation in the frequency bands of interest; sensing of the medium.
- Principles of direction finding and location: Single sensor techniques, including antenna arrays; multi-sensor techniques; jamming aspects.
- Impact of propagation on direction finding and location: Long waves; HF; VHF; UHF and above.
- Future requirements.

# Thème

Les systèmes de défense modernes doivent être dotés de fonctions de surveillance très complètes. L'un des éléments clé d'un système de surveillance est la capacité de détecter et de localiser les signaux radio. Les cahiers des charges de tels équipements couvrent non seulement la surveillance globale mais aussi les zones locales en situation de déploiement rapide des forces. Tout le spectre radio est concerné: les grandes ondes servent aux télécommunications stratégiques; la bande HF reste l'un des domaines de fréquences le plus utilisé dans tous les pays du monde pour les télécommunications globales et régionales; la gamme de fréquences allant de 50 MHz jusqu'au domaine millimétrique est largement utilisée pour les télécommunications et la surveillance.

Les effets du milieu de propagation jouent un rôle important dans l'application des techniques de radiolocalisation. Pour la haute fréquence et les fréquences inférieures l'ionosphère est le milieu principal; aux fréquences supérieures des effets environnementaux importants sont occasionnés par la réfraction atmosphérique. Sur toute la bande de fréquences radio, le bruit, aussi bien naturel qu'artificiel, joue souvent un rôle important dans les applications de goniométrie et de localisation.

Les techniques de radiolocalisation reposent sur l'angle d'arrivée du signal et la localisation de l'émetteur. Ces opérations peuvent être effectuées soit à partir d'un site unique (fixe ou mobile et même orbital) soit à partir d'un réseau de stations. Parmi les techniques qui ont été mises en oeuvre avec succès on peut citer la mesure de l'angle d'arrivée, le calcul de la différence du temps d'arrivée et la détermination de la direction Doppler d'arrivée des signaux. Les systèmes de radiolocalisation de la prochaine génération devront pouvoir employer des antennes réseau plus compactes avec une plus grande mobilité, fournir des solutions plus adéquates, fusionner des données provenant de différents systèmes, et tenir compte du milieu de propagation de façon automatique et transparente pour l'utilisateur. Ils feront appel à des systèmes experts pour l'analyse des données et le dialogue avec l'utilisateur.

## Sujets Traités

- Généralités: Propagation dans les bandes de fréquence intéressantes, télédétection du milieu.
- Les principes de la goniométrie et de la radiolocalisation: Techniques à un seul capteur, y compris les antennes réseau; techniques multi-capteurs; aspects brouillage.
- L'impact de la propagation sur la goniométrie et la télédétection: Les grandes ondes; HF; VHF; UHF et au-delà.
- Les besoins futurs.



# Electromagnetic Wave Propagation Panel

**Chairman:** Dr J.H. Richter  
Naval Command, Control and Ocean  
Surveillance Center  
RDT&E Division, Code 54  
San Diego, CA 92152-5000  
United States

**Deputy Chairman:** Dr D.H. Hohn, apl. Prof.  
FGAN  
Forschungsinstitut für Optik  
Schloss Kressbach  
D-7400 Tübingen  
Germany

## TECHNICAL PROGRAMME COMMITTEE

### Chairmen

Dr J.H. Richter  
NCCOSC  
RDT&E Division, Code 54  
San Diego, CA 92152-5000  
United States

Mr R.B. Rose  
NCCOSC  
Code 542  
San Diego, CA 92152-5000  
United States

### Members

Prof. C. Goutelard  
Directeur L.E.T.T.  
Université Paris-Sud  
Bâtiment 214  
F-91405 Orsay Cedex  
France

Prof. T.B. Jones  
Department of Physics  
University of Leicester  
University Road  
Leicester LE1 7RH  
United Kingdom

Dr R. Klemm  
FGAN-FFM  
Neuenährer Strasse 20  
D-5307 Wachtberg 7  
Germany

Dr D. Rother  
SEL LS E  
Lorenzstrasse 10  
D-7000 Stuttgart 40  
Germany

Prof. G. Tacconi  
DIBE  
Università di Genova  
Via Opera Pia 11a  
I-16145 Genova  
Italy

Prof. Dr C. Toker  
Department of Electrical Engineering  
Middle East Technical University  
Ankara  
Turkey

Dr D. Yavuz  
PO Box 174  
SHAPE  
N-2501 CD The Hague  
The Netherlands

## PANEL EXECUTIVE

Lt-Col. R. Cariglia

**Mail from Europe:**  
AGARD—OTAN  
Attn: EPP Executive  
7, rue Ancelle  
F-92200 Neuilly-sur-Seine  
France

**Mail from US and Canada:**  
AGARD—NATO  
Attn: EPP Executive  
Unit 21551  
APO AE 09777

Tel: 33(1)47 38 57 68  
Telex: 610176 (France)  
Telefax: 33(1)47 38 57 99

## Preface

The "Radio Location Techniques" symposium is an excellent example of AGARD's unique role and capability. Since the topic is primarily of interest in military applications and includes sensitive areas, it would be difficult to handle in an unrestricted forum. The subject matter includes complex propagation scenarios, novel signal processing techniques, sophisticated hardware, and extensive measurement programs. A good interaction between university, industry and both civilian and military members of defence establishments is necessary to address the interdisciplinary aspects of the subject matter. AGARD is in the unique position to facilitate such an interaction and has demonstrated it again with this highly successful symposium.

The symposium was held in London, U.K. from 1st—4th June 1992 and included seven sessions. Two sessions addressed propagation aspects, one radio location techniques, two radiolocation measurements and systems and one signal processing aspects. One session was classified and the proceedings will be published in a separate classified volume. All together, 39 papers were presented during the symposium.

Gratefully acknowledged are the cooperation and assistance of the Technical Programme Committee — Professor Goutelard, Professor Jones, Dr Klemm, Dr Rother, Professor Tacconi, Professor Toker and Dr Yavuz. In addition to members of the Technical Programme Committee, Dr Norbury and Professor Yeh served as session Chairs.

Appreciation is furthermore expressed to the AGARD staff led by Lieutenant Colonel Canglia and to Mr Jean-Philippe Prouteau for the excellent planning and organization. Very special thanks go to the local hosts, Mrs S. Martin, Dr P. Cannon and Dr K. Craig provided a memorable setting and superb arrangements which resulted in an excellent symposium.

J.H. Richter and R.B. Rose

# Contents

	Page
<b>Theme/Thème</b>	iii
<b>Electromagnetic Wave Propagation Panel</b>	iv
<b>Preface</b>	v
	<b>Reference</b>
 <b>SESSION I – PROPAGATION EFFECTS ON RADIOLOCATION (A)</b> Session Chairman: Dr J.H. Richter	
<b>Lateral Deviation of VLF Radio Waves due to Diffraction and Scattering from Coast-Lines, Mountain Ranges and Polar Ice Caps</b> by J.R. Wait	1
<b>Propagation Characteristics of the Ionospheric Transmission Window Relating to Long Wave Radio Location Issues</b> by P.A. Kossey and E.A. Lewis	2
<b>The Effects of Multipath Scatter from Wind Driven Gravity Waves on Lines of Bearing Serving as Direction Finders</b> by R.H. Ott	3
<b>Propagation Effects on HF Direction Finding</b> by G.H. Millman	4
 <b>SESSION II – PROPAGATION EFFECTS ON RADIOLOCATION (B)</b> Session Chairman: Dr J.R. Norbury	
<b>Nouvelle Méthode de Modélisation Mésoscale de l'Ionosphere à Partir d'un Site Unique. Application à la Radiolocalisation HF</b> (New Method of Modelling the Mesoscale Structure of the Ionosphere with a Single Sensor. Impact on HF Radiolocation) par J.L. Mokrzycki, L. Barthes, J. Caratori et C. Goutelard	5
<b>The Spatial and Temporal Characteristics of High Frequency Auroral Backscatter</b> by D.S. Choi, B. Weijers, R.J. Norris and N.B. Myers	6
<b>The Longitudinal Occurrence of Equatorial F Layer Irregularities</b> by J. Aarons	7
<b>Irregular Media Effects on Radiowave Signals used in Navigation and Positioning Systems</b> by E. Bahar	8
<b>Influence de la Propagation sur la Goniométrie Haute Fréquence et la Localisation à Station Unique</b> par V. Baltazart, L. Bertel et R. Fleury	9
 <b>SESSION III – RADIOLOCATION TECHNIQUES</b> Session Chairman: Prof. K.C. Yeh	
<b>A Current Assessment of Single Site Locating Technology</b> by R.B. Rose	10
<b>Paper 11 withdrawn</b>	

	Reference
<b>Digital Processing for Positioning with One Satellite</b> by A. Marguinaud	40
<b>A New Passive SSL Technique for HF Radiolocation</b> by R.L. Johnson, Q.R. Black and A.G. Sonsteby	12
<b>Procedures for Determining the Location of an HF Radio Transmitter</b> by L.F. McNamara	13
<b>High Resolution Sampled Aperture Array Direction Finding at HF</b> by T.N.R. Coyne	14
<b>Detection of Active Emitters using Triangulation and Trilateration Techniques: Theory and Practice</b> by A.M. Dean	15
<b>SESSION IV – RADIOLOCATION MEASUREMENTS AND SYSTEMS (A)</b> Session Chairman: Dr D. Yavuz	
<b>Radiolocation of a Satellite-Borne LOVHF Beacon</b> by R.B. Rose	16
<b>Systematic Bearing Errors of HF Signals Observed over a North-South Propagation Path</b> by T.B. Jones, E.M. Warrington and J.E. Perry	17
<b>Large Bearing Errors Observed at a High Latitude DF Site</b> by T.B. Jones and E.M. Warrington	18
<b>Measurements of the Direction of Arrival of an Oblique Chirp Sounder Signal</b> by E.M. Warrington, T.B. Jones and P. Hamadyk	19
<b>Bearing Determination and Bearing Quality Indication from a Goniometric HF DF System</b> by E.M. Warrington and T.B. Jones	20
<b>Les Séquences W.G. – Séquences Binaires Quasi Parfaites</b> (The W.G. Sequences – Quasi Perfect Binary Sequences) par C. Goutelard	42
<b>SESSION V – RADIOLOCATION MEASUREMENTS AND SYSTEMS (B)</b> Session Chairman: Prof. Dr C. Toker	
<b>Paper 21 withdrawn</b>	
<b>Radiolocation in the Lower ELF Frequency Band</b> by C.P. Burke and D. Llanwyn Jones	41
<b>Angle of Arrival Characteristics of Ionospheric Skywave Signals</b> by D.M. Haines and B.W. Reinisch	22
<b>Utilisation du Réseau de Réception du Sondeur à Rétrodiffusion de l'Île Losquet en Radiogoniométrie</b> par J.Y. Le Saout, F. Gauthier, N. Ruelle et R. Fleury	23
<b>Test Results of the Advanced Translator Processing System</b> by A.K. Brown, W. Sward and P. Brown	24
<b>Communication Interference/Jamming and Propagation Analysis System and its Application to Radio Location</b> by H. Kuzucu	25
<b>Airborne System for Detection and Location of Radio Interference Sources</b> by B. Audone and A. Pastore	26

## SESSION VI — CLASSIFIED SESSION

Session Chairman: Mr R.B. Rose

- L'Evolution des Techniques de Goniométrie pour la Surveillance Radio dans les Bandes HF et VUHF — de l'Interférométrie aux Techniques de Goniométrie à Haute Résolution** 27\*  
(Evolution of Direction Finding Techniques in Radio Surveillance for the HF and VUHF Band — from the Interferometry to the High Resolution Direction Finding Technique)  
par G. Multédo
- Etude sur la Localisation d'Emetteurs à Partir d'une Station Aéroportée** 28\*  
(Study on Emitter Location from an Airborne Platform)  
par C. Vaudet
- Paper 29 withdrawn**
- Direction Finding and Location Principles for Short Time Signal Interception and Reconnaissance** 30\*  
by E. Koelble
- An Emitter Location System for Airborne Application** 31\*  
by R. Schmelzer
- Goniomètre pour Hélicoptères Utilisant la Diffraction par la Voilure tournante** 32\*  
(Direction Finder on Board Helicopters using Rotor Blades Scattering)  
par F. Christophe et A. Sarremejean
- Influence des Réflexions Structurelles sur le Traitement des Algorithmes de Localisation à Bord des Avions Tactiques** 33\*  
(Influence of Structural Reflections on Threat Localization Algorithms on Board Tactical Aircraft)  
par P. Fossier et C. Maillard
- Localization System Based on Data Merging of Images and Inertia** 34  
by J.Y. Cartoux, G. Sella, L. Agranier, T. Péchoux and G. Grenier

## SESSION VII — SIGNAL PROCESSING FOR RADIOLOCATION

Session Chairman: Prof. C. Goutelard

- Théorie Fractale des Grands Réseaux d'Antennes Lacunaires** 35  
(Fractal Theory for Large Lacunar Antenna Arrays)  
par C. Goutelard
- Localisation d'Emissions à Evasion de Fréquence et Rafales par Capteur à Analyse Séquentielle en Balayage Rapide** 36  
par D. Josset
- Radio Location through High Resolution Eigenstructure Processing Techniques that Yield Accurate Multipath AOA and Differential Time Delay Signal Estimates** 37  
by L.M. Lewandowski and K.A. Struckman
- Multiple Emitter Direction Finding from a Circular Array by Single Snapshot Processing** 38  
by Y. Tanik and A. (Tuncay) Koç
- Automatic Feature Extraction and Localization using Data Fusion of Radar and Infrared Images** 39  
by S. Houzelle and G. Giraudon

\* Published in classified volume CP-528 (Supplement).

# LATERAL DEVIATION OF VLF RADIO WAVES DUE TO DIFFRACTION AND SCATTERING FROM COAST-LINES, MOUNTAIN RANGES AND POLAR ICE CAPS

James R. Wait  
Consultant  
2210 East Waverly  
Tucson Arizona  
85719 3848 USA

## SUMMARY\*

This paper deals with a physical theory of radio wave propagation in the earth-ionosphere waveguide when allowance is made for lateral non-uniformities in the lower boundary. A combined hybrid ray and mode analysis is adopted which reduces to the now conventional formulation when the lower boundary is uniform. For obvious reasons, the mathematical details will not be presented but the optical interpretation is given and the physical concepts illustrated in a hopefully meaningful fashion. A case of particular interest is when the global propagation of the VLF signal arrives from both the direct path as well as a sometimes very strong resonant scatter from a continental mountain range 1000 km or more off the great circle path. Other examples occur when the signal actually diffracts around convex coast-lines and continental margins such as the Antarctic ice cap. In such cases, the direct path signal over the highly lossy medium is highly attenuated and may be out of contention.

In the simplest theory of "horizontal bending" of VLF waveguide modes, the conversion of modes, of one order to another, is ignored. But, by employing mode match techniques, the conversion of modes of order  $m$  to modes of order  $n$  (reflected) and to modes of order  $p$  (transmitted) can be estimated. Both a scalar and a vector version of the formulation are outlined. Actually the scalar version is really more appropriate for dealing with acoustic ducts with inhomogeneous walls when the horizontal direction of propagation is oblique to the junction planes. Strictly speaking the full blown vector problem must be dealt with in the electromagnetic case if proper allowance is to be made for the inevitable coupling between the TM(transverse magnetic) and TE(transverse magnetic) modes.

## 1 INTRODUCTION

To understand some of the complexities of bearing errors in VLF radio wave transmission, it is desirable to consider an idealized model which can later be employed as a basis for discussing more realistic situations where the physics is very similar. The situation is shown in Fig. 1a, b, and c below. A source (e.g. a vertical electric dipole) is located at  $(0, y_0, 0)$  on the earth's surface  $z=0$ . The upper reflecting boundary is at  $z=h$  and this surface is assumed to be laterally uniform and characterized by a surface

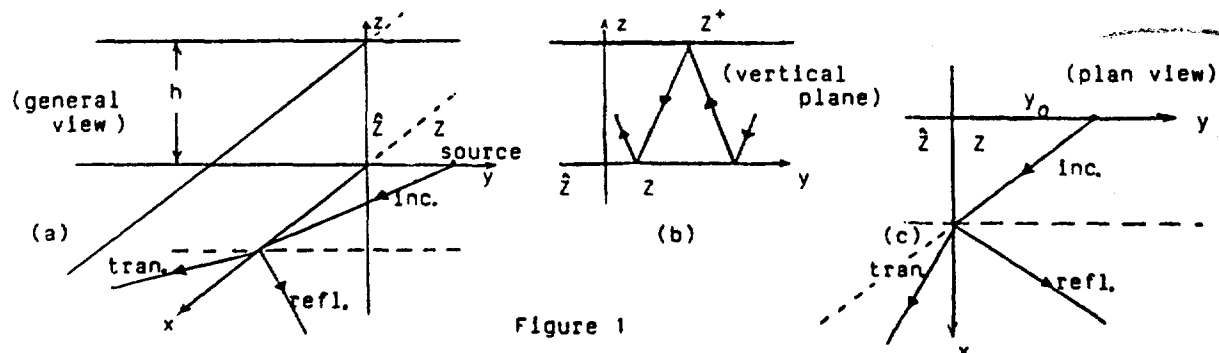


Figure 1

impedance  $Z^+$ . The lower boundary, at  $z=0$ , is a two part surface whose impedance is  $Z$  for  $y > 0$  and is  $\hat{Z}$  for  $y < 0$ . A more explicit description of these boundary conditions is given below.

\* The first paragraph of the summary refers to the oral presentation of the paper at the AGARD Symposium in London, June 1992. The second paragraph is pertinent to the theory.

In a qualitative sense, we now consider that the source dipole emanates a waveguide mode of order  $m$  in the semi-infinite region  $y > 0$ . It is "reflected" at the junction plane  $y = 0$  into a spectrum of modes of order  $n$  including the specular case  $n = m$ . The transmitted modes of order  $p$  propagate into the region  $y < 0$ . Again, if there was no mode conversion  $p = m$  but, even then, the direction of the wave propagation will be modified. It is this modification, appropriately defined below, which we will call the bearing error. In general, converted modes (i.e. where  $p \neq m$ ) will also suffer bearing errors and they may be very significant.

Some very interesting situations may arise at highly oblique incidence where the reflected or scattered modes may become significant while the transmitted modes could become very small in magnitude and highly damped in the direction of transmission. Such a situation is analogous to critical reflection at a dielectric interface when the waves are going from a dense to a less-dense region.

It is very important to recognize that we have a double-barreled problem to contend with. In plan view, as indicated in Fig. 1b, we are dealing with a ray picture but, in the vertical plane, as indicated in Fig. 1c, we have more of a waveguide mode picture. In other words, an individual "vertical mode" is treated as a ray in the horizontal plane. This concept is not new and, indeed, seismologists and underwater acousticians have often exploited the technique. But, surprisingly, radiowave people do not usually treat such mixed path problems in this context. (see references 1 to 4).

## 2. A SCALAR MODEL

In our formulation we will adopt a scalar viewpoint, even though the 3d problem is intrinsically vector (as we indicate in Section 5). In this context, we will deal with a complex field function  $\psi(x, y, z)$  which satisfies the Helmholtz equation  $(\nabla^2 + k^2)\psi = 0$  (1) in the region  $0 < z < h$  where  $k = w/c = 2\pi/\text{free space wavelength}$  and  $w$  is the angular frequency. The implied time factor is  $\exp(+j\omega t)$  here and in what follows. We can identify  $\psi$  as the vertical electric field for VLF waves but in other related wave problems it can have other meanings. (e.g. the excess pressure in acoustic wave guides).

The impedance type boundary conditions that we invoke on the horizontal boundaries are listed as follows:

- i  $\partial\psi/\partial z = -jk\Delta^+\psi$  at  $z = h$  for  $-\infty < y < \infty$  (2)
- ii  $\partial\psi/\partial z = jk\Delta\psi$  at  $z = 0$  for  $y > 0$  (3)
- iii  $\partial\psi/\partial z = jk\hat{\Delta}\psi$  at  $z = 0$  for  $y < 0$  (4)

where we have adopted normalized surface impedances defined as follows:  $\Delta^+ = Z^+/\bar{R}_0$ ,  $\Delta = Z/\bar{R}_0$ , and  $\hat{\Delta} = Z^+/\bar{R}_0$  where  $\bar{R}_0 = 120\pi$  ohms. These normalized impedances are also relevant to analogous acoustic models.

Now, of course, we recognize that a source dipole at  $(0, y_0, 0)$  will excite a spectrum of discrete modes but it is sufficient to consider one such mode of order  $m$  because, in the final analysis, superposition holds and the appropriate sum over  $m$  may be effected. Thus we choose our incident wave function, for a far distant source, to be of the form:

$$\psi^{inc} = \exp(-jkS_m \sin \theta_m x - jkS_m \cos \theta_m y) G_m(z) \quad (5)$$

where unimportant or slowly varying factors are omitted and where  $G_m(z)$  is the height gain function defined below. Also we should note that  $\theta_m$  is the real angle of incidence in the horizontal  $(x, y)$  plane for the incident waveguide mode of order  $m$ .

The appropriate eigen or characteristic modes in the vertical plane are obtained from

$$[(C - \Delta)/(C + \Delta)][(C - \Delta^+)/(C + \Delta^+)] \exp(-2jk h C) = \exp(-j2\pi n) = 1 \quad (6)$$

The solutions are  $C = C_m$  which yield  $S_m = (1 - C_m^2)^{1/2}$ . The physical constraints on the solution are that  $\text{Re}.S_m > 0$  and  $\text{Im}.S_m < 0$ . Also we normalize the height gain

functions such that  $G_m(0) = 1$  which follows convention. Now from (1), we see that

$$G_m(z) = [\exp(+jkC_m z) + R_m \exp(-jkC_m z)] / (1 + R_m) \quad (7), \quad \text{where } R_m = (C_m - \Delta) / (C_m + \Delta) \quad (8)$$

Here  $R_m$  can be identified as the reflection coefficient at the ground surface  $z = 0$  for the complex eigen angle,  $\arccos C_m$ , for the mode of order  $m$ . We may readily verify that (1) is satisfied and that the boundary conditions (2) and (3) are met.

For such an incident waveguide mode, as defined by (5), we now write down the expected form of the reflected modes of order  $n$ . The field function of  $\psi_n^{\text{refl}}$  is of the form

$$\psi_n^{\text{refl}} = R_{n,m} \exp(-jkS_n \sin \theta_n x - jkS_n \cos \theta_n y) G_n(z) \quad (9); \quad \text{where } S_n = (1 - C_n^2)^{1/2} \quad \text{and}$$

$G_n$  is a solution of (6) for the  $n$ 'th mode.  $R_{n,m}$  is the reflection coefficient for the incident mode of order  $m$  as it is reflected and/or converted into mode  $n$ .

To deal with the transmitted modes of order  $p$ , we are led to write, in analogy to (9):

$$\psi_p^{\text{trans}} = T_{p,m} \exp(-jk\hat{S}_p \sin \hat{\theta}_p x + jk\hat{S}_p \cos \hat{\theta}_p y) \hat{G}_p(z) \quad (10),$$

where now the vertical eigen modes are obtained from the solution  $C = C_p$  of

$$[(C - \hat{\Delta}) / (C + \hat{\Delta})] [(C - \Delta^*) / (C + \Delta^*)] \exp(-2jkCh) = \exp(-2jnp) = 1 \quad (11)$$

Then  $\hat{S}_p = (1 - \hat{C}_p^2)^{1/2}$  where again  $\text{Re. } \hat{S}_p > 0$  and  $\text{Im. } \hat{S}_p < 0$ . The height function

is now  $\hat{G}_p(z) = [\exp(jk\hat{C}_p z) - \hat{R}_p \exp(-jk\hat{C}_p z)] / (1 + \hat{R}_p) \quad (12), \quad \text{where}$

$$\hat{R}_p = (\hat{C}_p - \hat{\Delta}) / (\hat{C}_p + \hat{\Delta}) \quad (13). \quad \text{Then } T_{p,m} \text{ is the transmission coefficient}$$

for the incident mode of order  $m$  as it is converted into mode of order  $p$ .

Now at this point, the only assumption we need make is to say the transition, from the surface impedance  $Z$  for  $y > 0$  to the surface impedance  $\hat{Z}$  for  $y < 0$ , is invariant with respect to  $x$ . Thus  $\partial \psi_p^{\text{trans}} / \partial x = \partial \psi_n^{\text{refl}} / \partial x = \partial \psi_m^{\text{inc}} / \partial x = -jg_m \quad (14)$ , where  $g_m$  is a fixed parameter. (14) is a generalized form of Snell's law for the 3D problem.

The following properties are a consequence of (14): 1) Since  $\theta_m$  is real,  $\theta_n$  (angle of reflection) is generally complex except for  $n = m$  and for the unlikely situation that  $\arg. S_n = \arg. S_m$  when  $n \neq m$ , 2)  $\hat{\theta}_p$  (angle of transmittance) is generally complex even for  $n = m$  except, in the trivial case, where there is no surface impedance contrast.

We can now construct the desired forms for the modal representations for the total fields for an incident mode of order  $m$ . Thus, for  $y > 0$ , we may write for the total field:

$$\psi = \psi^{\text{inc}} + \psi^{\text{refl}} = A_m [\exp(u_m y) G_m(z) + \sum_n R_{n,m} \exp(-u_n y) G_n(z)] \exp(-jg_m x) \quad (15)$$

while, for  $y < 0$ ,

$$\psi = \psi^{\text{trans}} = A_m \sum_p T_{p,m} \exp(v_p y) \hat{G}_p(z) \exp(-jg_m x) \quad (16)$$

Here  $A_m$  is a coefficient specifying the strength of the incident  $m$ 'th order mode. The

other key parameters are:  $u_m = (g_m^2 - k^2 S_m^2)^{1/2} = j(k^2 S_m^2 - g_m^2)^{1/2} \quad (17),$

$$u_n = (g_m^2 - k^2 S_n^2)^{1/2} = j(k^2 S_n^2 - g_m^2)^{1/2} \quad (18), \quad v_p = (g_m^2 - k^2 \hat{S}_p^2)^{1/2} = j(k^2 \hat{S}_p^2 - g_m^2)^{1/2} \quad (19)$$

and where, as indicated by (14),  $g_m = k S_m \sin \theta_m$ .

### 3. PROPERTIES OF REFLECTED AND TRANSMITTED MODES

A procedure to determine the reflected and transmitted mode strengths is described later.

But, first of all, it is desirable to discuss the nature of the modes with regard to attenuation, phase velocity, and refraction. For example let us deal with the reflected modes. An individual mode will have the form  $\psi_n^{\text{refl}} = \exp(-jg_m x) \exp(-u_n y) \quad (21)$ , where we have omitted the multiplicative factors which do not depend on  $x$  and  $y$ . Now let us write  $g_m = \rho_m - j q_m \quad (22)$ , and  $u_n = \alpha_n + j \beta_n \quad (23)$ , where  $\rho_m, q_m, \alpha_n$ , and  $\beta_n$  are typically real and positive. Thus note that

$$\psi_n^{\text{refl}} = \exp[-q_m x - \alpha_n y] \exp[-j(\rho_m x + \beta_n y)] \quad (24). \quad \text{Clearly planes of}$$



of constant phase are determined by  $p_m x + \theta_n y = \text{const.}$  (25), or  $p_m = \theta_n dy/dx = 0$  (26). The situation is illustrated in Fig. 2. The direction of the wave fronts, for the modes of order  $n$ , make an angle  $\delta_{n,m}$  with the  $y$  axis. indeed  $\tan \delta_{n,m} = -dy/dx = p_m / \theta_n$  is an explicit formula to deduce the real angle  $\delta_{n,m}$  which is relevant to any phase sensitive receiving system. Of course, in the special case  $n=m$ , we see that  $\delta_{m,m} = \theta_m$ . Now looking at (24), we see that the planes of constant amplitude are determined by  $q_m x + \alpha_n y = \text{const.}$  (28), or  $q_m + \alpha_n dy/dx = 0$  (29). The angle of the normal,  $\delta'_{n,m}$ , is given by  $\tan \delta'_{n,m} = -dy/dx = q_m / \alpha_n$  (30). Except in the special case  $p_m / q_m = \theta_n / \alpha_n$  (31), (or  $\arg. S_m = \arg. S_n$ ), it is clear that  $\delta'_{n,m} \neq \delta_{n,m}$ . Therefore it is evident that the planes of constant phase and planes of constant amplitude do not coincide for the converted modes except where noted above.

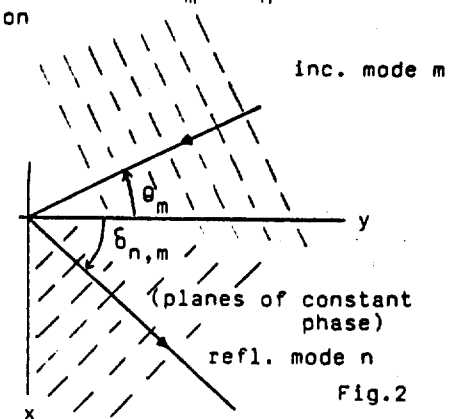


Fig. 2

The situation for the transmitted modes, in the region  $y < 0$  is similar. Such a mode of order  $p$  will have an  $x$  and  $y$  dependence as follows:  $\psi_p^{\text{trans}} = \exp(-j\theta_p x) \exp(v_p y)$  (32) where again we have omitted factors that do not depend on  $x$  and  $y$ . We now set  $v_p = \hat{\alpha}_p + j\hat{\beta}_p$  (33), where  $\hat{\alpha}_p$  and  $\hat{\beta}_p$  are real. Then  $\psi_p^{\text{trans}} = \exp(-q_p x + \hat{\alpha}_p y) \exp[-j(p_m x - \hat{\beta}_p y)]$  (34).

The planes of constant phase are now determined by  $p_m x - \hat{\beta}_p y = \text{const.}$  (35), or  $p_m - \hat{\beta}_p dy/dx = 0$  (36). The situation is illustrated in fig. 3. The refraction angle  $\hat{\delta}_{p,m}$ , describing the direction of the wave normal, is determined from  $\tan \hat{\delta}_{p,m} = dy/dx = p_m / \hat{\beta}_p$  (37). Except in the trivial case of no surface impedance contrast (i.e.  $Z = \hat{Z}$ ), the angle  $\hat{\delta}_{p,m} \neq \theta_m$  even in the case where  $p = m$ . Looking at (34), we see that planes of constant amplitude are determined from  $q_m x - \hat{\alpha}_p y = \text{const.}$  (38) or  $q_m - \hat{\alpha}_p dy/dx = 0$ . The refraction angle for planes of constant amplitude is given by  $\tan \hat{\delta}'_{p,m} = q_m / \hat{\alpha}_p$  (40). In general  $\hat{\delta}'_{p,m} \neq \hat{\delta}_{p,m}$  except again in the special case where the waveguides are lossless such that  $\arg. S_m = \arg. \hat{S}_p = 0$ . One might now ask: how does (40) compare with Fermat's principle? To answer this question, we note that  $v_p = jk[\hat{S}_p'^2 - S_m'^2 \sin^2 \theta_m]^{1/2}$  (41) can be approximated, for low order modes, by  $v_p = jk[(\hat{S}_p')^2 - (S_m' \sin \theta_m)^2]^{1/2}$  (42), where  $S_m' = \text{Re. } S_m$  and  $\hat{S}_p' = \text{Re. } S_p$ . Then, from (37),  $\tan \hat{\delta}_{p,m} = [S_m' \sin \theta_m] / [(\hat{S}_p')^2 - (S_m' \sin \theta_m)^2]^{1/2}$  (44), which trigonometrically is equivalent to  $\hat{S}_p' \sin \theta_{p,m} = S_m' \sin \theta_m$  (45). This latter form of Snell's Law becomes exact when  $\hat{S}_p$  and  $S_m$  are real.

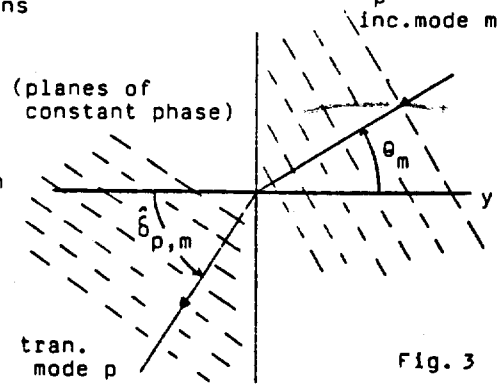


Fig. 3

#### 4. MODE MATCHING

We now develop a procedure to determine the reflection and transmission coefficients  $R_{n,m}$  and  $T_{p,m}$ , respectively. The method is based on mode matching. The scheme is to match  $\psi$  and  $\partial\psi/\partial y$  at the junction plane  $y=0$  over the interval  $0 < z < h$ . Employing (15) and (16), we arrive at the following:

$$G_m(z) + \sum_n R_{n,m} G_n(z) = \sum_p T_{p,m} \hat{G}_p(z) \quad (46),$$

$$\text{and} \quad u_m G_m(z) - \sum_n u_n R_{n,m} G_n(z) = \sum_p v_p T_{p,m} \hat{G}_p(z) \quad (47).$$

To deal with this system of equations, we can exploit the orthogonality of the height gain functions. In the present context, we multiply both sides of (46) and (47) by  $G_q(z)$  and integrate over  $z$  from 0 to  $h$ . Then we obtain:

$$(\Lambda_q)^{-1}(d_{m,q} + R_{q,m}) = \sum_p T_{p,m} c_{p,q} \quad \text{and} \quad (\Lambda_q)^{-1}(u_m d_{m,q} - u_q R_{q,m}) = \sum_p v_p T_{p,m} c_{p,q} \quad (48) \quad (49)$$

where  $d_{m,q} = 1$  for  $m=q$ ,  $= 0$  for  $m \neq q$ . In arriving at (48) and (49) we have exploited the following orthogonality property :

$$\text{where } \Lambda_q = (h/2) \left[ \int_0^h [G(z)]^2 dz \right]^{-1} \quad (51) \quad (2/h) \int_0^h G_p(z) G_q(z) dz = 1/\Lambda_q \text{ for } p=q \quad (50) \\ = 0 \text{ for } p \neq q$$

is a modal excitation parameter. Also we have defined  $c_{p,q} = (2/h) \int_0^h \hat{G}_p(z) G_q(z) dz$  (52) as a mode coupling characterization.

Now we can eliminate  $R_{q,m}$  from the pair (48) and (49) to yield the single set

$$(\Lambda_q)^{-1}(u_q + u_m) d_{m,q} = \sum_p (u_q + v_p) T_{p,m} c_{p,q} \quad (53)$$

To solve this system, we truncate the series over  $p$  to say  $N$  terms. Then we take  $q = 1, 2, 3, \dots, N$  and solve the finite system for  $T_{p,m}$  for the designated  $m$ 'th incident mode. To illustrate the concept, let us take  $m = 1$  and  $N = 2$  which is a rather drastic truncation. Then (53) reduces to  $(u_1 + v_1) T_{1,1} c_{1,1} + (u_1 + v_2) T_{2,1} c_{1,2} = 2u_1/\Lambda_1$  (54)

and  $(u_2 + v_1) T_{1,1} c_{2,1} + (u_2 + v_2) T_{2,1} c_{2,2} = 0$  (55) Consequently, to within

$$\text{first order } T_{2,1} = T_{1,1} [(u_2 + v_1)/(u_2 + v_2)] (c_{2,1}/c_{2,2}) \quad (56)$$

$$\text{and } T_{1,1} = (\Lambda_1)^{-1} [2u_1/(u_1 + v_1)] (1/c_{1,1}) = (\text{roughly}) 2u_1/(u_1 + v_1) \quad (57)$$

As we see the coupling coefficient  $c_{q,p}$  plays a key role. It can be expressed simply, as we shall show, following a somewhat classical route forged by the "Master" (i.e. A.N. Sommerfeld). First of all, we note that, over the interval  $0 < z < h$ , we can write :

$$(d^2/dz^2 + k^2 c_q^2) G_q(z) = 0 \quad (58) \quad \text{and} \quad (d^2/dz^2 + k^2 \hat{c}_p^2) \hat{G}_p(z) = 0 \quad (59)$$

Now we multiply (58) by  $\hat{G}_p$  and (59) by  $G_q$  and then subtract one of the resulting equations from the other to yield

$$\hat{G}_p d^2/dz^2 G_q - G_q d^2/dz^2 \hat{G}_p + k^2 (c_q^2 - \hat{c}_p^2) = 0 \quad (60)$$

But (60) is equivalent to

$$d/dz [\hat{G}_p dG_q/dz - G_q d\hat{G}_p/dz] = -k^2 (c_q^2 - \hat{c}_p^2) G_q \hat{G}_p \quad (61)$$

Then, on integrating both sides of (61) over the range of  $z$  from 0 to  $h$ , we get

$$[\hat{G}_p dG_q/dz - G_q d\hat{G}_p/dz] \Big|_0^h = -k^2 (c_q^2 - \hat{c}_p^2) (h/2) c_{q,p} \quad (62)$$

$$\text{But } dG_q/dz = jk\Delta \text{ at } z=0, \quad d\hat{G}_p/dz = jk\hat{\Delta} \text{ at } z=0 \quad (63), (64)$$

$$= -jk\Delta^* G_q \text{ at } z=h, \quad = -jk\hat{\Delta}^* \hat{G}_p \text{ at } z=h \quad (65), (66)$$

Thus the LHS (left hand side) of (62) is simply  $-jk(\hat{\Delta} - \Delta)$  so that the coupling coefficient is given by :

$$c_{p,q} = (2/(jkh)) (\hat{\Delta} - \Delta) / (c_q^2 - \hat{c}_p^2) \quad (67)$$

$$\text{Perhaps it is useful to note that } c_{2,1}/c_{1,1} = (c_1^2 - \hat{c}_1^2) / (c_1^2 - \hat{c}_2^2) \quad (68)$$

which is an alternative form showing that the mode conversion is most important when the modes 1 and 2 become nearly coincident (i.e. near degeneracy).

It is now useful to return to (46) and (47) in an attempt to deal explicitly with the reflection /conversion coefficient  $R_{n,m}$ . To this end, we play the same orthogonality game and begin by multiplying both sides of (46) and (47) by  $\hat{G}_s(z)$  and integrating over  $z$  from 0 to  $h$ . Then we obtain the system :

$$c_{s,m} + \sum_n R_{n,m} c_{s,n} = T_{s,m} / \hat{\Lambda}_s \quad (69)$$

$$\text{where } \hat{\Delta}_s = (h/2) \left[ \int_0^h [\hat{G}_s(z)]^2 dz \right]^{-1} \quad (71) \quad u_m c_{s,m} - \sum_n u_n R_{n,m} c_{s,n} = v_s T_{s,m} / \hat{\Lambda}_s \quad (70)$$

and

$$\text{and } c_{s,n} = (2/h) \int_0^h G_n(z) \hat{G}_s(z) dz = (2/(jkh)) (\hat{\Delta} - \Delta) / (c_n^2 - \hat{c}_s^2) \quad (72)$$

On eliminating  $T_{s,m}$ , from (69) and (70), we arrive at the single set

$$(v_s - u_m) + \sum_n (v_s + u_n) R_{n,m} c_{s,n} / c_{s,m} = 0 \quad (73)$$

To obtain numerical values for  $R_{n,m}$  we truncate the summation over  $n$  to  $N$  terms and then let  $s = 1, 2, 3, \dots, N$ . The resulting matrix of the coefficients is then inverted. In the rather trivial case where mode conversion is neglected (i.e.  $N = 1$ ), we obtain:

$R_{m,m} = (u_m - v_m) / (u_m + v_m)$  (74) which is very simple. In fact (74) can be interpreted as the reflection coefficient at an interface between two half spaces of refractive indices  $S_m$  and  $\hat{S}_m$ . But, in general, we need to deal with (73) for some specified number of terms. In this case, it is useful to note that the ratio of the coupling coefficients is given by  $c_{s,n} / c_{s,m} = (c_m^2 - \hat{c}_s^2) / (c_n^2 - \hat{c}_s^2)$  (75) which is consistent with (67).

## 6 GRADUAL TRANSITION OR NON ABRUPT BOUNDARY

In the foregoing analysis, we have idealized the junction as an abrupt change of the conductivity of the lower boundary (e.g. a straight flat lying coastline). But, in many cases, the junction is tapered in some fashion. For example, in the analogous day/night transition, the width of the earth-ionosphere waveguide will change gradually. Similarly an irregular coastline will often appear as a tapered transition. Here we will indicate if such a transition or tapered junction will modify any of our conclusions about bearing errors. The model we consider is a modification of the geometry considered earlier in Fig. 1. The situation we deal with now is shown in Fig. 4 where, in fact, rather than having a sudden change of properties at  $y = 0$ , we have a transition that extends from  $y = -d$  to  $+d$ . The incident mode of order  $m$ , for  $y > d$ , is again specified to be  $\psi^{inc} = \psi_0 \exp(-jg_m x) \exp(u_m y) G_m(z)$  where  $u_m = jkC_m$  where  $C_m$  satisfies the mode eqn. (6). Now, in the region  $y < -d$ , the transmitted mode of order  $m$  will still have the form  $\psi^{trans} = \psi_0 T \exp(-jg_m x) \exp(v_m y) \hat{G}_m(z)$  in the absence of any mode conversion.  $G_m(z)$  and  $\hat{G}_m(z)$  are the appropriate height gain functions in the regions  $y > d$  and  $y < -d$ , respectively. Now in the transition region,  $d > y > -d$ , we will argue that the solution will have the form:  $\psi(y) = \exp(-jg_m x) Y(y) G_m(y, z)$  (74) where the  $y$  dependent height gain function

is given by  $G_m(y, z) = [\exp[jk C(y) z] + R_m \exp[-jk C(y) z]] / (1 + R_m)$  (75)

where  $R_m = [C_m(y) - \Delta(y)] / [C_m(y) + \Delta(y)]$  (76)

and where  $\Delta(y)$  is the  $y$ -dependent (normalized) surface impedance at  $z = 0$  for  $-d < y < d$ .  $C_m(y)$  then satisfies the  $y$ -dependent mode equation such as (6) with  $\Delta$  replaced by  $\Delta(y)$ . Now since  $\psi$  satisfies the Helmholtz equation  $(\nabla^2 + k^2)\psi = 0$ , (77), in the space  $0 < z < h$ , it is clear that  $[d^2/dy^2 + k^2 q^2(y)] Y(y) = 0$ , (78), where  $q(y) = j(k^2 S_m^2(y) - g_m^2)^{1/2}$  where  $S_m(y) = [1 - C_m^2(y)]^{1/2}$ . The appropriate WKB solution of (78) is:

$$Y(y) = [1/kq(y)]^{1/2} \exp[jk \int_0^y q(y) dy] \quad (79) \quad \text{apart from a constant. Now this}$$

solution form must match the incident wave as  $y \rightarrow +\infty$  and the transmitted wave as  $y \rightarrow -\infty$ . In these two limits  $q(y) \rightarrow C_m$  and  $q(y) \rightarrow \hat{C}_m$ , respectively. These conditions are met when

$$T = (C_m / \hat{C}_m)^{1/2} \exp[-jk \int_{-d}^d q(y) dy] \exp[jk (C_m - \hat{C}_m) 2d] \quad (80)$$

Of course, it is not permitted to let  $d$  approach zero because the WKB solution is only valid if  $q(y)$  or  $S_m(y)$  is a slowly varying function of  $y$ . In such a case, mode conversion would also be small. But an interesting observation can be made about (80), if  $g_m > k \hat{S}_m$ ,  $v_m$  in (73) becomes mainly real. Then we can expect that there will be critical reflection back into the region  $y > 0$  from whence the incident wave came.

For example, it is possible a value of  $y = y_c$ , in the transition, will occur when  $g(y) = 0$  (or there will be a minimum). Of course the problem is analogous to the reflection of obliquely incident radio waves (in the vertical plane) at the ionosphere where the electron density increases with height. Some of the relevant theory on phase integral methods can be taken over directly (e.g. see Ref.3). In this case we may deduce that for the region  $y > d$ , the critically reflected field will be given by  $\psi^{refl} = \psi_0 j \exp[-jk_c(y - 2y_c)]$  (81) where the  $\pi/2$  phase lead is associated with the effective caustic surface at  $y = y_c$ . The situation as described above is depicted in fig. 4 where two typical situations are shown; in the one case, the incident mode has a horizontal angle of propagation which is close to perpendicular with the  $y$  axis. Then the "ray" is continuously bent within the transition zone but it eventually exits at  $y = -d$  and goes on its way into region  $y < -d$ . On the other hand, the highly oblique ray enters the transition zone but it is continually refracted back towards the region  $y > d$ .

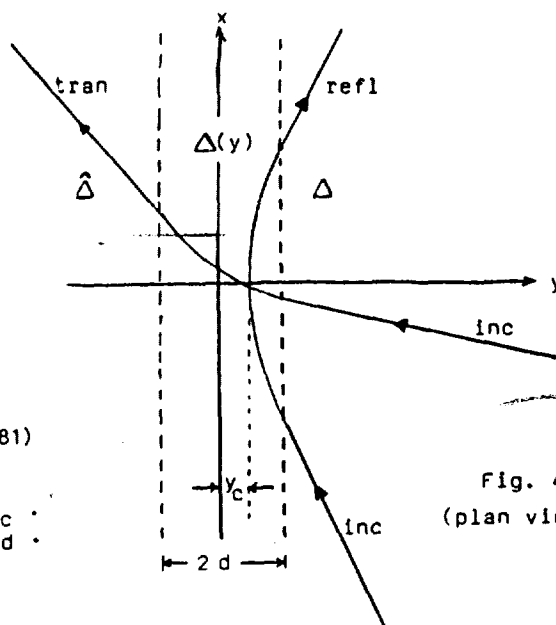


Fig. 4  
(plan view)

[Illustrating a penetrating mode and a critically reflected mode]

Perhaps it is important to note that with such a transition zone, the properties of the modes in the uniform regions  $y < -d$  and  $y > +d$  are precisely the same as we discussed earlier for an abrupt boundary - everything else being equal. Of course, the relative amount of mode conversion will be a strong function of the width or the extent of the transition zone. But it is safe to say that the bearing errors for a given mode should not be affected by the width or taper of the transition zone.

## 7. ALLOWANCE FOR TE-TM COUPLING

A more rigorous analysis of the oblique incidence problem is to allow for coupling between the TE (transverse electric) and the TM (transverse magnetic) waves. We shall outline the procedure for dealing with this situation. The waveguide model is the same as illustrated in Fig. 1. In the region  $0 < z < h$ , the fields can be derived most conveniently from the Debye potentials  $U$  and  $V$  (Ref.2). Thus we may write

$$E_x = \partial^2 U / \partial x \partial z - j\mu_0 \partial V / \partial y \quad (81), \quad E_y = \partial^2 U / \partial y \partial z + j\mu_0 \partial V / \partial x \quad (82),$$

$$H_x = \partial^2 V / \partial x \partial z + j\epsilon_0 \omega \partial U / \partial y \quad (83), \quad H_y = \partial^2 V / \partial y \partial z - j\epsilon_0 \omega \partial U / \partial x \quad (84),$$

$$E_z = (k^2 + \partial^2 / \partial z^2) U \quad (85) \text{---and} \quad H_z = (k^2 + \partial^2 / \partial z^2) V \quad (86)$$

Here we associate the  $U$  with the TM or E Modes while  $V$  is associated with the TE or the H Modes. Essentially, in our previous scalarized or acoustic-like model, we ignored the fields derived from the TE or H modes (i.e. we assumed  $H_z = 0$ ). It might be mentioned here in passing that in dealing with a realistic earth-ionosphere model, where the earth's magnetic field renders the ionosphere anisotropic, there is also coupling between the TM and TE modes even for complete lateral non-uniformity of the wave guide (see Refs. 2 & 3). That is a complication not considered here.

To proceed with our analysis, we construct our solutions for  $U$  and  $V$  that satisfy the Helmholtz equation:  $(\nabla^2 + k^2) U$  (or  $V$ ) = 0 (87). The incident wave is a pure TM or E mode, of order  $m$ , as assumed earlier. Thus  $E_z^{inc} = \exp(-jg_m x + u_m y) G_m(z)$  (88)

where  $q_m = k S_m \sin \theta_m$  and  $u_m = (q_m^2 - \kappa^2 S_m^2)^{1/2} = j(\kappa^2 S_m^2 - q_m^2)^{1/2}$  being consistent with our earlier definitions. Again  $\theta_m$ , the (horizontal) angle of incidence for mode  $m$ , is real. Also, as before,  $S_m = (1 - C_m^2)^{1/2}$  where  $C_m$  satisfies the mode equation (6) as applicable to the region  $y > 0$ . It is now clear that  $\partial/\partial x$ , in equations (81) to (84), can be replaced by  $-jg_m$  with the same justification as in the scalarized version. The boundary conditions, on the horizontal surfaces, are now stated to be:

$$\begin{aligned} E_x &= -\tilde{n}_0 \Delta H_y \Big|_{y>0, z=0} & E_x &= -\tilde{n}_0 \hat{\Delta} H_y \Big|_{y<0, z=0} & E_x &= +\tilde{n}_0 \Delta H_y \Big|_{x>y>-\infty, z=h} \\ E_y &= +\tilde{n}_0 \Delta H_x \Big|_{y>0, z=0} & E_y &= +\tilde{n}_0 \hat{\Delta} H_x \Big|_{y<0, z=0} & E_y &= -\tilde{n}_0 \Delta H_x \Big|_{x>y>-\infty, z=h} \end{aligned} \quad (89), (90), (91)$$

We now write the appropriate forms of the Debye potentials, for the incident mode of order  $m$ , in the compact forms, where summation over the repeated indices is understood:

$$U_m = \exp(u_m y) G_m(z) + R_{n,m}^e \exp(-u_m y) G_n(z) \quad (92), \quad V_m = R_{s,m}^h \exp(-u_s y) F_s(z) \quad (93); \quad y > 0$$

$$U_m = T_{p,m}^e \exp(v_p y) \hat{G}_p(z), \quad (94), \quad V_m = T_{q,m}^h \exp(v_q y) \hat{F}_q(z) \quad (95); \quad y < 0.$$

where, as before,  $v_p = j(\kappa^2 \hat{S}_p^2 - q_p^2)^{1/2}$  but also  $v_q = j(\kappa^2 \hat{S}_q^2 - q_q^2)^{1/2}$  where

$\hat{S}_q$  (lower case  $s$ ) =  $(1 - \hat{C}_q^2)^{1/2}$  in terms of  $\hat{C}_q$  which is the solution of the TE mode discussed below. Similarly,  $u_n = j(\kappa^2 c_n^2 - q_n^2)^{1/2}$  but now, also  $u_s = j(\kappa^2 s_s^2 - q_s^2)^{1/2}$

where  $s_s = (1 - c_s^2)^{1/2}$  in terms of  $c_s$  which is the solution of the corresponding TE mode equation. The TE type height gain functions are then very conveniently written in the form

$$F_s(z) = \exp(jk c_s z) + r_s \exp(-jk c_s z) \quad (96) \quad \text{where} \quad r_s = (\Delta c_s - 1)/(\Delta c_s + 1)$$

$$\hat{F}_q(z) = \exp(jk \hat{C}_q z) + \hat{r}_q \exp(-jk \hat{C}_q z) \quad (97) \quad \text{where} \quad \hat{r}_q = (\hat{\Delta} \hat{C}_q - 1)/(\hat{\Delta} \hat{C}_q + 1)$$

Now  $c_s$  is a solution of the TE mode equation:  $r_s r^+ \exp(-j2kch) = \exp(-j2\pi s) = 1$  (98)

and  $\hat{C}_q$  is a solution a similar mode equation:  $\hat{r}_q r^+ \exp(-j2kch) = \exp(-j2\pi q) = 1$  (99)

where  $r^+ = (\Delta^+ c - 1)/(\Delta^+ c + 1)$ . Then the corresponding TM mode equations are identical to those employed in the earlier scalarized formulation.

The remaining task is to match the tangential fields  $E_x$ ,  $E_z$ ,  $H_x$  and  $H_z$  in the junction. This step leads to the following simultaneous equations for the unknown coefficients:

$$-jg_m G_m'(z) + R_{n,m}^e G_n'(z) + j\mu w R_{s,m}^h u_s F_s(z) = -jg_m T_{p,m}^e \hat{G}_p'(z) - j\mu w T_{q,m}^h v_q \hat{F}_q(z), \quad (100)$$

$$-jg_m R_{s,m}^h F_s'(z) + j\mu w [\mu_m G_m(z) - R_{n,m}^e u_n G_n(z)] = -jg_m T_{q,m}^h \hat{G}_q'(z) - j\mu w T_{p,m}^e v_p \hat{G}_p(z), \quad (101)$$

$$C_m^2 G_m(z) + C_n^2 R_{n,m}^e G_n(z) = \hat{C}_p^2 T_{p,m}^e \hat{G}_p(z), \quad (102)$$

$$C_s^2 R_{s,m}^h F_s(z) = \hat{C}_q^2 T_{q,m}^h \hat{F}_q(z) \quad (103) \quad \text{where} \quad G_n'(z) = dG_n(z)/dz, \text{ etc.}$$

These four systems of equations must hold

and  $F_s'(z) = dF_s(z)/dz$ , etc.

over the interval  $0 < z < h$ . No further attempt

here will be made to solve them. But a number of general conclusions can be drawn.

For perpendicular incidence, (i.e.  $g_m = 0$ ), the TE modes become completely decoupled and the results are identical to the scalarized model. Also it may be noted that for the low order modes, the derivatives of the height gain functions are relatively small and the coupling between TM and TE modes is weak. Finally it is to be noted that at highly oblique incidence, the modes may be critically reflected in which case the coupling tends to be small.

### 3. REFERENCES

- 1) J.B. Keller and J.S. Papadakis, (editors), "Wave Propagation and Underwater Acoustics" Springer Verlag, Heidelberg, 1977 (see Chap. 3)
- 2) J.R. Wait, "Wave Propagation Theory", Pergamon Press, New York, 1981 (see Chap. 20)
- 3) K. G. Budden, "Propagation of Radio Waves", Cambridge University Press, 1985
- 4) J.R. Wait, "VLF reflection from mountain ranges", Jour. Atmos. Terr. Phys. **54**, 193-196, 1992

# ANNOTATED BIBLIOGRAPHY OF VLF RADIO WAVE PROPAGATION

Monographs, Book Chapters, Review Articles, and Journal Special Issues:

- 1-1 J.R. Wait (editor), VLF Symposium Papers, Proc. Inst. Radio Eng. vol. 45, 739-810, no. 6, 1957 ( contains 14 papers selected from a symposium held in Boulder in Jan. 1957)
- 1-2 K.G. Budden, "The Wave-guide Mode Theory of Wave Propagation", 325 pp. Prentice-Hall, Englewood Cliffs, 1961 ( based on a course of lectures give in June 1957 , at Pacific Naval Lab., Esquimalt, BC, Canada where the similarity of sound wave transmission in the ocean and long radio waves is stressed. The emphasis throughout the book is on understanding the mathematical methods rather than on their immediate practical use.)
- 1-3 W.T. Blackband, "Propagation of Radio Waves at Frequencies Below 300 kc/s", 477 pp. Macmillan Co., New York, 1964 ( an edited collection of articles from a NATO/Agard conference held in Munich in 1962. )
- 1-4 A.D. Watt, "VLF Radio Engineering" 701 pp. Pergamon Press, Oxford 1967 ( a masterful treatment of all aspects of very low frequency radio waves including very understandable descriptions of the basic concepts of the propagation phenomena and system design)
- 1-5 W.T. Blackband (chairman), "M.F., L.F., and V.L.F. Radio Propagation", Conference publication No. 36, 350 pp., Inst. of Elect. Eng. (U.K.) 1967 ( includes good survey articles on the : the lower ionosphere, waveguide mode theory, atmospheric VLF phenomena in the magnetosphere, and phase stability)
- 1-6 H. Volland, "Die Ausbreitung Langer Wellen", 232 pp. Friedr. Vieweg & Sohn, Braunschweig, 1968 ( a self contained exposition of the propagation theory at VLF and LF )
- 1-7 J.R. Wait, "Electromagnetic Waves in Stratified Media" pp. 608 pp., Pergamon Press, Oxford, 2nd. Ed. 1970 ( 50% of the book is on VLF propagation theory)
- 1-8 Ya. L. Alpert and D.S. Fligel, "Propagation of ELF and VLF Waves Near the Earth" 171 pp. Plenum Publ. Co. New York, 1970 ( a translation of the Soviet monograph published in 1967 and contains some important material on the effect of the earth's H field)
- 1-9 J. Galejs, "Terrestrial Propagation of Long Electromagnetic Waves" pp. 362, Pergamon Press, Oxford, 1972 ( a good treatment of the whole subject and includes good summaries of the author's fundamental contributions covering 10 years prior to his tragic death)
- 1-10 J.R. Wait (editor), Special Issue on "Extremely Low Frequency (ELF) Communications" IEEE Trans. on Communications, vol. 22, 353-588, 1974 ( includes material relevant to VLF)
- 1-11 V.G. Pezrodny, A.P. Nickolaenko and V.G. Sinitsin, Radio propagation in natural waveguides, Jour. of Atmos. & Terr. Phys., vol. 39, 661-688, 1977 ( a good survey of recent theoretical work by Soviet investigators on waves in natural channels such as the earth-ionosphere waveguide with some interesting material on wall roughness)
- 1-12 H. Volland (editor), "Handbook of Atmospheric" Vol. 2, pp. 327, CRC Press, Boca Raton, Florida, 1982 ( contains several chapters on propagation of 'sferics' which are signals originating from lightning )
- 1-13 F.J. Kelly, "ELF/VLF/LF Propagation and System Design" Report 9028, 1987, US Naval Research Lab., Wash. D.C. (updates much of the material in Ref. 1-4 above and includes some excellent discussions on modulation techniques, airborne antennas, disturbances and noise)

## Mainly Theoretical Papers with Emphasis on Modal Representations:

- 2-1 J.R. Wait, The mode theory of VLF ionospheric propagation for finite ground conductivity, Proc. Inst. radio Eng. vol. 45, 760-767, 1957 ( showed that the hitherto ignored finite ground conductivity could significantly influence VLF attenuation)
- 2-2 J.R. Wait, An extension to the mode theory of VLF ionospheric propagation, Jour. Geophys. Res. vol. 63, 125-135, 1958 ( used a two layer ionosphere to show modified frequency dependence of the attenuation rate and phase velocity of the dominant modes)
- 2-3 J.R. Wait, A diffraction theory for LF sky wave propagation, Jour. geophys. Res. vol. 66, 1713-1724, 1961 ( a wave theory for the case where the first hop skywave propagates into the shadow). An additional note, ibid 1725-1729, 1961
- 2-4 J.R. Wait, On the propagation of VLF and ELF radio waves when the ionosphere is not sharply bounded, Jour. Res. NBS vol. 66D(Radio Prop.), 53-61, 1962
- 2-5 J.R. Wait and K.P. Spies, Influence of earth curvature and the terrestrial magnetic field on VLF propagation, Jour. Geophys. Res. vol. 65, 2325-2331, 1960
- 2-7 K.G. Budden, The influence of the earth's magnetic field on radio propagation by waveguide modes, Proc. Royal Soc., A, vol. 265, 538-553, 1962
- 2-8 K.G. Budden, The ionosphere as a whispering gallery, ibid., 554-569, 1962
- 2-9 J.R. Wait, Excitation of modes at VLF in the earth-ionosphere waveguide, Jour. Geophys. Res. vol. 67, 3823-3828, 1962 ( develops theory for the 'earth-detached modes')
- 2-10 J. Galejs, Excitation of V.L.F. and E.L.F. radio waves by a horizontal magnetic dipole, Jour. Res. NBS, vol. 65D (Radio Prop.) 305-311, 1961
- 2-11 J. Galejs, Terrestrial ELF propagation, in "Natural Electromagnetic Phenomena Below 30 kc/s" (edited by D.F. Bleil), plenum press, pp. 205-258, 1964
- 2-12 L.A. Berry and M.E. Chrisman, The path integrals of LF/VLF wave hop theory, Radio

Science (Jour. Res. NBS, vol. 69D, 1469-1480, 1965 ( extends ref. 2-3 above)

2-13 J. Galejs, Propagation of very low freq. waves below an anisotropic stratified ionosphere with a transverse static magnetic field, Radio Science, vol. 2, 557-574, 1967

2-14 J. Galejs, Propagation of ELF and VLF waves below an anisotropic ionosphere with a dipping static magnetic field, Jour. Geophys. Res. vol. 73, 339-352, 1968

2-15 R.A. Pappert, A numerical study of VLF mode structure and polarization below an anisotropic ionosphere, Radio Science, vol. 3, 219-232, 1968 (uses a cylindrical model and an application of a conformal transformation to incorporate the earth's magnetic field in the full wave mode calculations - following crucial suggestions of J.H. Richter)

2-16 J.R. Wait, The whispering gallery nature of the earth-ionosphere waveguides at VLF, Trans. IEEE, vol. AP-15, 580-581, 1967 ( Correction, vol. AP-16, 147, 1968)

2-17 J.R. Wait, Mode conversion and refraction effects in the earth-ionosphere waveguide for VLF radio waves, Jour. Geophys. Res. vol. 73, 3537-3548, 1968 (considers the non- abrupt day/night transition and bearing errors for oblique transmission)

2-18 J.R. Wait, Note on mode conversion at VLF in the earth-ionosphere waveguide, *ibid.* vol. 73, 5801-5804, 1968 (considers a rudimentary model for lateral diffraction)

2-19 J.R. Wait and K.P. Spies, On the calculation of mode conversion at a graded height change in the earth-ionosphere waveguide at VLF, Radio Science, vol. 3, 787-791, 1968

2-20 J.R. Wait, On mode conversion of VLF radio waves at a land/sea boundary, Trans. IEEE vol. AP-17, 216-217, 1969 ( develops an approximate mode matching procedure and shows conversion between dominant modes is appreciable for land of poor conductivity)

2-21 J.R. Wait, Factorization method applied to EM wave propagation in a curved waveguide with non-uniform walls, Radio Science, vol. 5, 1059-1068, 1970 ( shows that earlier approximate methods for calculating mode conversion at land/sea boundaries are valid)

2-22 R. Pappert and R.R. Smith, Orthogonality of VLF height gains in the earth ionosphere waveguide, Radio Science, vol. 7, 275-278, 1972 ( exploits the concept of the adjoint guide)

2-23 R.A. Pappert and F.P. Snyder, Some results of a mode-conversion program for VLF, Radio Science, vol. 7, 913-923, 1972 (allows for the vertical inhomogeneity of the ionosphere as well as its anisotropy)

2-24 K.G. Budden, The critical coupling of modes in a tapered earth-ionosphere waveguide, Math. Proc. Camb. Phil. Soc., vol. 77, 567-580, 1975 (examines the situation where two modes in the twilight region have coincident propagation constants)

2-25 R. Smith, Mode conversion coefficients in the earth-ionosphere wave guide for VLF propagation below a horizontally stratified, anisotropic ionosphere, Jour. Atmos. Terr. Phys., vol. 39, 539-543, 1977 (employs a factorization method to justify the neglect of the reflected modes when computing the forward mode conversion)

2-26 R.A. Pappert and J.A. Ferguson, VLF/LF mode conversion model calculations for air to air transmissions in the earth-ionosphere waveguide, Radio Science, vol. 21, 551-558, 1986 ( discusses the simplification of the full wave determination of the biorthogonal set of height gains which, in effect, neglects the gain functions above some height)

2-27 J.R. Wait, VLF Mode conversion for ionospheric depressions, Radio Science, vol. 26, 1261-1265, 1991 ( allowance is made for the depression to be laterally depressed from the great circle path ) (Note in eqn. 24,  $S_{1,2}$  should be  $S_{2,1}$ )

2-28 B.J. Rask and A.D. Gorman, On caustics associated with horizontal rays and vertical modes, Wave Motion, vol. 14, 321-332, 1991 ( considers the caustics that may occur in the ray theory ansatz for treating slowly varying media in the lateral direction)

2-29 J.R. Wait, Diffraction of VLF radio waves by polar ice caps, Electronics Letters, vol. 27, 1030-1031, 1991 ( an example where mode theory is used in the vertical direction and generalized ray theory in the lateral directions ) ( note eqn. 8 reads  $\partial D / \partial y - qD = 0$  )

2-30 J.R. Wait, Reflection of VLF radio waves at a junction in the earth-ionosphere waveguide, IEEE Trans. vol. EMC 34, 4-8, 1992 ( a scalar model is employed to deal with the reflected modes which are quite small except at highly grazing incidence ) ( note that  $C_{2,1}$  in eqn. 24 should be replaced by  $C_{2,1} R_{2,1}$ , also  $C_{2,2}$  in eqn. 25 should be replaced by  $C_{2,2} R_{2,1}$  ) ( In eqns. 5 and 6, the factors  $\exp(jk_m z)$  and  $\exp(jk_p z)$  were missed on RHS's)

Mainly Experimental Papers relevant to VLF Propagation Including Phase Variations:

3-1 E. Bahar and J.R. Wait, Propagation in a model terrestrial waveguide of non-uniform height: Theory and Experiment, Radio Science, Jour. Res. NBS, vol. 69D, 1445-1467, 1965 ( uses a generalized reciprocity theorem to account for reflected waves and follows this up with a laboratory microwave modelling experiment )

3-2 J. Frisius, On the determination of VLF propagation parameters by field strength measurements over medium distances, Radio Science (New Series), vol. 1, 513-522, 1966 (a very careful series of measurements at 16 kc/s (kHz) from GBR, Rugby, at distances from 350 to 800 km employing simple reflection coefficient parametric model following Wait)

3-3 E. Bahar, Propagation of VLF radio waves in a model earth-ionosphere waveguide of arbitrary height and finite surface impedance boundary: theory and experiment, Radio Science, vol. 1, 925-938, 1966 ( day-night transition and localized depressions modelled)

- 3-4 J. L. Rhoads and W. L. Garner, An investigation of the modal interference of VLF radio waves, *Radio Science*, vol. 2, 539-546, 1967 ( observations of field strength measured in Washington, DC, for transmissions from Hawaii at frequencies between 16.6 and 26.1 kHz showing good agreement for predicted modal interference in Daytime using Wait's model)
- 3-5 G. L. Alpert, D. S. Fligel and G. A. Michailova, The propagation of atmospherics in the earth-ionosphere waveguide, *Jour. Atmos. Terr. Phys.*, vol. 29, 29-42, 1967 (shows some experimental data for transient waveforms for lightning source signals at ranges from 300 to 1000 km indicating amazing agreement with calculated wave shapes using mode theory)
- 3-6 R. B. Orlov and G. V. Azarnin, Main Laws of propagation of VLF signals in the earth-ionosphere waveguide, "Diffraction and Propagation Radiowave Problems", vol. 10, 3-107, 1970, Leningrad State University, ( an exhaustive survey of experimental results for the frequency range from 5 to 30 kHz showing critical comparisons between theory and observations using both natural and artificial sources)
- 3-7 R. Barr, The propagation of ELF and VLF radio waves beneath an inhomogeneous anisotropic ionosphere, *Jour. Atmos. Terr. Phys.* vol. 33, 343-353, 1971 (computed values of attenuation are in good agreement with mode theory except for the highly attenuated region between 1 and 4 kHz but but observed phase velocities show broader discrepancies)
- 3-8 R. Barr, The effect of the earth's magnetic field on the propagation of ELF and VLF radio waves, *Jour. Atmos. Terr. Phys.* vol. 33, 1577-1583, 1971 ( uses a realistic inhomogeneous anisotropic ionosphere model and shows that azimuthal dependence is very similar to Wait's (1962) earlier heuristic estimates )
- 3-9 E. C. Field and C. Greifinger, Transpolar propagation of long radio waves, *Jour. Geo-Phys. Res.* vol. 77, 1264-1278, 1972 ( shows calculations that predict gross distortions of VLF modes when propagating over ice cap with strong amplitude decreases and phase advances which were in accord with their three - dimensional laboratory scale model)
- 3-10 S. Westerlund and F. H. Røder, VLF signals propagating over the Greenland ice sheet, *Jour. Atmos. Terr. Phys.* vol. 35, 1475-1491, 1973 ( calculations show phase velocity of 1st TM mode is higher over ice than over sea water while reverse is true for 2nd and 3rd modes but all three modes have higher attenuation which is consistent with observations)
- 3-11 D. L. Jones, ELF ionospheric radio propagation studies using natural sources, *IEEE Trans.* vol. Com. 22, 477-484, 1974 ( while mainly relevant to ELF, shows some useful results for observed slow tails of atmospheric signals which accompany the VLF signal)
- 3-12 W. Harth, Beschreibung von VLF-Atmospherics-Parametern mit dem Wait-und Walters-Modell, *Zeitschrift für Geophysik*, Band 38, 153-167, 1972 ( extends mode calculations to frequencies in the range 3 to 10 kHz and compares these results with experimentally derived attenuation rates and phase velocities showing good agreement )
- 3-13 Y. Muraoka, A new approach to mode conversion observed in mid-latitude VLF transmission, *Jour. Atmos. Terr. Phys.* vol. 44, 855-862, 1982 ( reports observations of amplitude and phase of 18.6 kHz transmission from NLK (Wash. State) to Japan for a three year period and deduces mode conversion coefficients for terminator)
- 3-14 R. Barr and P. O. Helm, Standing waves in VLF transmissions produced by geographical discontinuities, *Jour. Atmos. Terr. Phys.* 44, 967-971, 1982 ( shows clear evidence that standing waves occur, due to reflection from the Southern Alps of New Zealand, for a path from NWC in Australia such that a strong interference minimum occurs between the dominant modes just at the right place to make such a measurement very sensitive)
- 3-15 Y. Muraoka, Winter anomalous effects of mode conversion observed in mid-latitude VLF transmissions, *Jour. Geophys. Res.* vol. 88, 311-317, 1983 ( concludes that nighttime VLF reflection height decreases considerably during the D region winter anomaly)
- 3-16 A. Tolstoy and T. J. Rosenberg, Model predictions of subionospheric VLF signal perturbation from localized electron precipitation-induced ionization enhancement regions, *Jour. geophys. res.* vol. 91, 13,473-13,482, 1986 ( uses VLF mode theory to examine the effect of electron ionization enhancement including both mode conversion and changing ground conductivity along the path and such calculations support the whistler excitation mechanism)
- 3-17 U. S. Inan and D. L. Carpenter, Lightning-induced electron precipitation events observed at L = 2.4 as phase and amplitude perturbations on subionospheric VLF signals, *Jour. Geophys. res.* vol. 92, 3293-3303, 1987 ( phase Trimp events observed at Palmer Station Antarctica, were interpreted to extract information on precipitation flux density)
- 3-18 R. L. Dowden and C. D. D. Adams, Phase and amplitude perturbations on the NWC signal at Dunedin from lightning-induced electron precipitation, *Jour. Geophys. Res.* vol. 94, 497-503, 1987 ( uses Wait's 1964 theory for diffraction from a circular stalactite to interpret off great circle echoes)
- 3-19 W. L. Poulsen, T. F. Bell and U. S. Inan, 3D modelling of sub-ionospheric VLF propagation in the presence of localized D region perturbations associated with lightning, *Jour. Geophys. Res.* vol. 95, 2355-2366, 1990 ( generalizes Wait's 1964 model to include slow variations in both the transverse directions as well as along the great circle path)
- 3-20 W-Y Yip, U. S. Inan and R. E. Orville, On the spatial relationship between lightning discharges and propagation paths of perturbed subionospheric VLF/LF signals, *Jour. Geophys. res.* vol. 96, 249-258, 1991 ( concludes that the occurrence of Trimp events for L = 3 or so may be more controlled by magnetospheric conditions than source lightning distribution)



## DISCUSSION

### D. JONES

In an early interpretation of refraction effects, Neil Thomson used a very simple analysis based on refraction at an interface characterized by a refractive index computed in terms of the phase velocities on the two sides. How valid is this treatment?

### AUTHOR'S REPLY

I believe Thomson's method is valid for an individual mode. But he does not consider mode conversion nor does he deal with the magnitude of the mode conversion.

### J. BELROSE

Your reference to VLF attenuation rates for waveguide mode propagation across icecaps of 50 dB/1000 km is a bit of an exaggeration...or at least should be qualified. The "attenuation rate" is a strong function of frequency and reflection height, and is greatest at low VLF frequencies and low reflection heights. Thus, for practical VLF frequencies, it is greatest for the OMEGA frequencies (10 kHz) and lowest reflection heights (50 km, corresponding to a solar proton event). Perhaps such "attenuation rates" occur under such propagation conditions?

### AUTHOR'S REPLY

Yes. 50 dB is an exaggeration or upper limit. Perhaps such an opaque model of the Antarctic ice cap is appropriate in view of Barr's observations.

### R. OTT

In regard to the plot of reflection loss versus incidence angle, how do the results change for the higher order mode numbers? In general, how are the coupled equations for the transmission/reflection coefficients truncated and solved for the higher-order modes?

### AUTHOR'S REPLY

The lowest mode will be critically reflected in some cases, even when mode conversion is non-negligible. The higher order mode conversion is calculated by truncating the system of equations at successively higher numbers to secure convergence.

# **PROPAGATION CHARACTERISTICS OF THE IONOSPHERIC TRANSMISSION WINDOW RELATING TO LONG WAVE RADIO LOCATION ISSUES**

Paul A. Kossey  
Phillips Laboratory  
Geophysics Directorate (PL/GPI)  
Hanscom AFB, MA 01731 USA

Edward A. Lewis  
Systems Integration Engineering  
35 Bedford Street  
Lexington, MA 02173 USA

## **ABSTRACT**

Most applications of long radio waves (ELF/VLF/LF/MF) are ground-based and exploit the fact that such signals can propagate to great distances via reflections from the lower ionosphere. It is known however that, owing to the influence of the earth's magnetic field, long wave signals can penetrate through the ionosphere as well; at times, with relatively little loss, depending on ionospheric conditions and other propagation factors. This has prompted investigations of the long wave "ionospheric transmission window" as part of efforts to assess the feasibility of deploying long wave emitters in space for terrestrial applications and/or for exploiting, in space, signals emanating from ground-based long wave transmitters. This paper outlines results of theoretical and experimental investigations of the ionospheric transmission window over the frequency range from about 100 Hz to 500 kHz, with emphasis on directional issues associated with long wave penetration of the ionosphere.

## **1. INTRODUCTION**

For discussion purposes, it is convenient to refer to radio waves having frequencies below 3000 kHz as "long waves". Included are Extremely Low Frequencies (ELF), Very Low Frequencies (VLF), Low Frequencies

(LF), and Medium Frequencies (MF). Long waves propagate to distances well beyond line-of-sight via a number of different modes. These include propagation over the surface of the earth by diffraction modes (ELF/VLF/LF/MF); propagation by transmission-line type modes (ELF); and, propagation by earth-ionosphere waveguide modes (VLF/LF/MF).

ELF has an important application for military communications that require wave penetration beneath the surface of the ocean; the VLF/LF bands are used extensively for navigation and military communication; and, the MF band (from 535-1606 kHz) is used by the standard AM broadcast systems. Long waves are also used in basic ionospheric research, lightning studies, standard frequency and time distribution, geological studies, and minerals exploration. For almost all of these applications, long wave transmitters and/or receivers are deployed on, or relatively near, the surface of the earth.

It is well known that long radio waves can also penetrate through the densest regions of the ionosphere; i.e., through the F-max region, where the electron density is many orders of magnitude larger than that needed to completely reflect the waves if the ionosphere were a simple (unmagnetized) plasma. This transmission window is due to the geomagnetic field of the earth, which constrains the electron motion

produced by incident electromagnetic waves. The study of "whistlers" has shown that long waves from lightning strokes can penetrate the ionosphere, follow a geomagnetic field line, and return to the earth at a distant point. This type of hemisphere-to-hemisphere propagation occurs under special time-varying conditions which are still not fully understood. Related research has focussed on using long wave whistler-mode propagation to trigger electron precipitation events in a controlled manner. Such a capability could, ultimately, have interesting potential application in reducing the threat posed to space craft by charged particles in the radiation belts. Another potential application of whistler-mode propagation, a space-based ELF/VLF transmitter for communications applications, has been considered by the U.S. Air Force and Navy (most extensively in the late 1960's and early 1970's) but has never been carried past the conceptual stage owing, primarily, to the large technical uncertainties and potential high cost of deploying such a system.

Technological advances over the past decade and recent space-based experimental activities in this area, such as the Soviet Union's ACTIVE VLF satellite and NASA electron-beam (SEPAC) and long wire experiments (WISP), suggest that it might be useful to review some of the characteristics of the ionospheric transmission window, as it applies to long radio waves. More specifically, results of a number of experimental and computational programs conducted by the Phillips Laboratory of the United States Air Force will be described.

## 2. ROCKET INVESTIGATIONS OF VLF PENETRATION OF THE IONOSPHERE

While satellites such (e.g., Ref 1) offer many attractions (such as

coverage of very large geographic areas, repeated measurements at different times, etc.) for VLF ionospheric penetration studies, there is an important role for rocket probe experiments (e.g., Ref 2) because these can give vertical amplitude-height profile information needed for understanding and confirming the penetration mechanism. In the 1960's and 70's the U.S. Air Force conducted a number of rocket-probe experiments to explore the VLF transmission window of the ionosphere, measure the attenuations, and obtain other data of potential interest for engineering applications. The experimental approach is illustrated in Figure 1. Briefly, signals from distant VLF transmitters were monitored from the ground to about 600 km altitude with receiving systems contained in the nose cones of rockets flown from Eglin, Florida, Natal, Brazil, and Fort Churchill, Canada. These locations provided an opportunity to investigate the relative effect of the geomagnetic field on the long wave penetration properties, since the dip angle of the geomagnetic field ranged from near zero degrees at Natal, sixty degrees at Eglin, and eighty-four degrees at Ft. Churchill.

VLF receiving instrumentation was contained in the nose cones of sounding rockets, which reached apogees of about 600 km (well above the F-max of the ionosphere) during the flights. The payload included an antenna that was an electrostatically shielded air-core loop, shaped to conform to the fiberglass nose cone of the rocket. In flight, the axis of the rocket was nearly vertical so that the antenna sensed the horizontal component of the magnetic flux density of the signals being monitored. The antenna spun with the rocket at 6 rps typically, and information on wave polarization could be obtained by observing the resulting signal modulations. For a

linearly polarized transverse magnetic wave there were two maxima and two minima in each revolution, while for circularly polarized waves the signals had constant amplitudes. In the circularly polarized case the signal frequency differed from the wave frequency by an amount equal to the rotational frequency, being greater or less depending on the relative directions of rotation. With either polarization there were additional frequency shifts (Doppler) due to the component of rocket velocity in the wave propagation direction. When signal-to-noise conditions were favorable, the Doppler shift could be obtained by allowing for the relatively constant rotational shifts produced by the spinning rocket.

The base of the antenna contained a solid state receiver with individual tuned channels to monitor various VLF stations. On the ground the original RF signal monitored on the rocket was recovered from the telemetry signal and recorded on video tape. Another channel recorded the RF signal from a similar payload operated in a fixed position near the ground. On playback, the rocket and ground signals could be compared cycle by cycle to obtain Doppler information. It was also possible to measure delays in arrival times, especially when the signals consisted of dots and dashes. Since the percentage bandwidths of the VLF transmissions were generally small, the observed signal delays could be assumed to approximate mathematical group delays.

### 2.1. Signal Profiles for the Night Flight at Eglin, Florida

Three VLF stations, NSS (Annapolis, Maryland, 21.4 kHz), NAA (Cutler, Maine, 17.8 kHz), and NLK (Jim Creek, Washington, 18.6 kHz) were monitored during the Eglin night flight. The propagation paths were completely

dark during the flight except for a small portion of the NLK-Eglin path in the vicinity of the transmitter. The measured amplitude-height profiles observed during the flight are shown in Figure 2. For comparison purposes, the profiles have been normalized with respect to a reference value of unity on the ground.

Except for a few periods when the signal from NLK was lost in the noise, signals were received on the rocket from all three stations from the ground to altitudes of 500 km. The NSS and NAA signals began to decrease shortly after lift-off and passed through brief minima around 40 km altitude (100 km for NLK). After a broad minimum between 100-200 km (200-300 km for NLK) the NSS and NAA signals tended to recover partially, while the signal from NLK held relatively steady.

In the region between 400-500 km, the NLK, NSS, and NAA signals were 16, 6, and 12 dB, respectively, below their values at ground level. In this sequence it happened that the signals from the nearer stations penetrated the ionosphere better than those from the more distant ones, a fact which may be related to their effective angles of incidence, and possibly to the geomagnetic field conditions (this will be discussed in more detail later).

The amplitude variations from 250-500 km seen in the NAA profile might be due to a standing wave pattern set up by reflections from some still higher altitude or, by what may be more likely, the interference effects of the superposition of waves belonging to different modal systems in the earth-ionosphere waveguide. Since the multiple reflected waves associated with the higher order modes have steeper incidence angles and may penetrate the ionosphere better, it may be that their

influence on the shape of the wave amplitude profile is relatively more pronounced in and above the ionosphere than in the waveguide below.

#### 2.1.1. NAA Polarization Profile for the Eglin Night Flight

A measure of wave polarization was obtained from the "spin modulation" quantity  $(1 - A_{\min}/A_{\max})$ , where  $A_{\min}$  and  $A_{\max}$  are respectively, the minimum and maximum signal amplitudes observed during one revolution of the payload's loop antenna. Figure 3 shows the percent modulation-height profile for the signals from NAA. Near the ground the modulation was nearly 100 percent, indicating that the propagating wave in the earth-ionosphere waveguide was linearly polarized, as would be expected. Above 250 km the modulation was very slight, indicating circular polarization: i.e., the magnetic flux density vector was rotating, rather than oscillating. As seen by an observer looking in the direction of propagation, the rotation was clockwise, as would be expected for whistler propagation. The minima at 40, 110, and 190 km are roughly coincident with the signal amplitude minima shown earlier.

#### 2.1.2. NAA and NSS Signal Delays for the Eglin Night Flight

Figure 4 gives the variations of the NAA and NSS signal delays with altitude, obtained by monitoring the dots and dashes received on the rocket payload and comparing them to those received on the ground. The rocket signal delays lagged the signals on the ground by considerably more than the millisecond maximum delay that could be attributed to the delay over the telemetry link. The signal delays became noticeable around 250 km, and increased to 12 ms near apogee at 500 km.

#### 2.1.3. Interpretation of the Eglin Night Flight Data

The similarities of the signal delay profiles of NAA and NSS, and the basic features of the amplitude profiles shown earlier suggest that the fields observed in and above the ionosphere were due to power leaking off from the earth-ionosphere waveguide in a coherent manner from directly below the rocket probe. Figure 5(A) shows a representation of a model of this penetration mechanism consisting of three regions: (1) the earth-ionosphere waveguide, (2) a transition region in the lower ionosphere, and (3) a region of whistler wave propagation. It is postulated that the wave patterns in regions (2) and (3) are rigidly tied to the pattern in region (1), which carries the main energy flow, and that the whole wave pattern moves from left to right with the phase velocity of the wave guide. Since the phase velocity in region (3) can be very much less than in region (1), the equiphase surfaces are shown sharply bent back, being nearly horizontal around the altitude of maximum electron density.

In the rocket experiment, the vertical velocity of the receiving antenna relative to the wave produced a Doppler shift of a few cycles per second in the region of circular polarization (3). The Doppler shifts for the relatively strong NAA signal were converted to give the refractive index profile in Figure 5(B), based on the assumption that the rocket trajectory was approximately perpendicular to the wave fronts as suggested by the penetration model. The refractive index data was then converted to the electron densities shown in Figure 5(C). The solid curve depicts the derived electron densities, while the dotted one shows values up to the F-max region obtained with a conventional HF ionosounder operated at Cape Kennedy.

The agreement between the two results is very good.

The amount of tilt of the equiphase surfaces can be estimated using Snell's Law, the inclination being approximately  $\sin^{-1}(1/n)$ , where  $n$  is the index of refraction. For  $n = 36$ , from Figure 5(B), the angle of inclination is 1.6 degrees.

For a given electron density profile, the group delay is inversely proportional to the square root of frequency, so that, theoretically, the signal delay for NAA (17.8 kHz) should be about 10 percent larger than that for NSS (21.4 kHz). By closely comparing the signal delays at 500 km altitude, it was found that the observed signal delay for NAA actually was nearly 10 percent larger than that for NSS.

## 2.2. Signal Profiles for the Day Flight at Eglin

The amplitude-height profiles of the signals from NSS and NAA observed during the day flight at Eglin are given in Figure 6, where it is shown that the NSS signal was monitored to only about 90 km altitude, where the signal was 20 dB below the ground value and became lost in instrumentation noise. The profile is smoother than the one obtained on the night flight; e.g., instead of the minimum around 35 km (Figure 2), there is merely an inflection near the 25 km altitude. This is consistent with the notion that higher order modes are less prominent in the daytime.

The corresponding amplitude-height profile for NAA shows a rapid increase in signal shortly after cut-off. After reaching a -30 dB minimum at an altitude of 70 km, the signal recovered to a level around -20 dB which was maintained with only minor variations up to 500 km. This signal was about 8 dB less than

observed at night, due to collision losses in the daytime D-region of the ionosphere.

The signal from NAA was strong enough during the day flight to allow group delays to be estimated from its dots and dashes at a number of altitudes. At 520 km the delay was about 30 ms, about twice the corresponding nighttime delay, which would suggest that the average electron density at this altitude was approximately four times greater in the day. This is in reasonable agreement with estimates of total electron content.

## 2.3. Day Flight at Natal, Brazil

The wave penetration in the coherent model discussed earlier is controlled by the vertical component of the geomagnetic field, which permits the wave fronts to travel in an essentially vertical direction. At the magnetic equator the vertical field component is zero and the penetration mechanism fails, so that signal levels above a few hundred kilometers in the ionosphere should be vanishingly small. As a test of this "blind zone" hypothesis, signals from six VLF stations were monitored at the geomagnetic equator in a day rocket flight at Natal, Brazil. Except for minor fluctuations, all six decreased as the rocket rose, practically disappeared around 100 km, and were not observed again until after the rocket reached apogee (660 km) and returned to the 100 km altitude region. Figure 7 gives three of the amplitude-height profiles, NBA (Panama, 24 kHz), NSS, and NAA obtained during the flight. The NAA signal was strong enough on the ground to permit the conclusion that the signal above 100 km were at least 30 dB down. This contrasts with the results of the Eglin day flight, for which the NAA signal at some altitudes above 100 km was only down about 17 dB. The Natal results were consistent with the "blind zone"

hypothesis to within the dynamic range of the measurements.

Although intelligible signals were not detected above the noise above 90-100 km at Natal, sometimes curious "blips" appeared in the NAA signal channel. These had durations of about 10 ms, noticeably shorter than the 15 ms dots of the code being transmitted by the station, but longer than most sferics noise-bursts. No time correlation was found between the occurrence of a blip and any signals, signal changes, or noise bursts seen by the ground observation receiver. Yet, since no such blips appeared in any of the other signal channels which had comparable sensitivities and bandwidths, the possibility that the blips were somehow related to the NAA signal cannot be ruled out. Shown in Figure 7 is the profile of the combined NAA signal and noise drawn to a scale to include apogee. The blips mainly appeared in the altitude regions marked by A, B, and C.

If the blips observed in the experimental data were "real" the question arises as to how the NAA signals got to the high altitudes above Natal. It may be that they penetrated the ionosphere at a higher latitude, where there was a vertical geomagnetic field component, and were then somehow coupled to the horizontal geomagnetic field lines that the rocket probe from Natal intercepted. This requires further investigation.

#### 2.4. Day Flight at Fort Churchill, Canada

A rocket probe with a 660 km apogee was made at Fort Churchill, Canada, where the geomagnetic field is both stronger and more nearly vertical than at Eglin (see Table 1). The amplitude-height profiles for the three VLF stations monitored during the flight are shown in Figure 8. In

none of the cases was a signal definitely received and identified above about 70 km. This was unexpected based on the penetration model previously developed from the Eglin data and discussed in conjunction with Figure 5.

##### 2.4.1. Comparison of the Eglin and Ft. Churchill NAA Data

NAA is nearly equidistant from Ft. Churchill and Eglin, so that it is interesting to compare the daytime amplitude-height profiles that were obtained at those locations. Figure 9 shows both profiles to the same scale. Below 70 km the profiles are very similar; but above that altitude, the signal at Ft. Churchill failed to recover, with the penetrating signal, if any, being at least 20 dB weaker than at Eglin. In light of the penetration mechanism outlined earlier, this is surprising since the vertical component of the geomagnetic field at Ft. Churchill is even stronger than at Eglin. The difference might be due to higher D-region absorption at Ft. Churchill, or different geomagnetic field orientation with respect to the propagating signals from NAA, or a combination of both.

An aspect of the orientation issue, illustrated in Figure 10, can be cast as follows: for the northward directed geomagnetic field, is there an appreciable difference in the penetration losses for waves from NAA propagating to the south (toward Eglin, Florida) compared to those propagating to the north (toward Ft. Churchill)? The schematic suggests that the coupling of waves to the geomagnetic field might "smoother" (less lossy) for southward propagation. To investigate this, normalized 17.8 kHz wave amplitude-height profiles were computed using a nominal night-time electron density of the ionosphere. The problem was oversimplified in that the earth and

the ionosphere were taken to be plane surfaces, and the wave structure between them was considered to consist of up- and down-going plane waves, with incidence angles of 87.3 degrees, chosen to represent the first propagating mode with a magnetic-type ionospheric boundary. A full-wave computer program (Ref 3) was used to calculate the penetrating and reflecting waves, and the total horizontal magnetic field vectors (the major axes of the polarization ellipses) were then determined by vector addition. The resulting amplitude-height profile computed for north-to-south and south-to-north wave propagation directions, with respect to the geomagnetic field, are given in Figure 11. From these simple calculations, the southward propagating wave couples into the ionosphere with less loss than the northward propagating wave. However, it is not clear whether this effect alone is sufficient to account for the observed difference between the Eglin and Ft. Churchill NAA ionospheric penetration data. This problem, the points raised earlier concerning the penetration of higher order modes, and the "blips" observed in the NAA data at the geomagnetic equator all suggest that further ionospheric penetration experiments would be necessary to help resolve such issues. The Ft. Churchill VLF penetration experiments, however, completed the U.S. Air Force's experimental research in this area.

### 3. COMPUTATIONS OF LONG WAVE IONOSPHERIC PENETRATION

A number of computational studies performed in conjunction with the rocket-probe experiments have bearing on some of the penetration issues outlined above. These illustrate clearly the variability of the penetration of long radio waves through the ionosphere in terms of a variety of factors, including wave frequency, ionospheric conditions,

geomagnetic field conditions, wave incidence angle on the ionosphere, and direction of wave propagation with respect to the geomagnetic field.

#### 3.1: Ionospheric Penetration Losses for Normal/Disturbed Conditions

Figure 12 shows computations by Booker, Crain, and Field (Ref 4) of daytime long wave penetration losses for normal (ambient) and disturbed ionospheric conditions, for waves propagating parallel to the geomagnetic field. As such, the results represent a "best" case (as discussed later). Under normal conditions, the daytime penetration losses are below about 30 dB for frequencies below about 20 kHz. The "disturbed" case represents a situation in which electron densities in the ionosphere below 70 km were 2-3 orders of magnitude higher than under normal conditions; a condition that sometimes occurs for relatively long periods in the polar cap following large eruptions on the sun (so-called solar proton events, or SPE's). Except for the very lowest frequencies (below about 100 Hz), the increased absorption in the lower is effectively prohibitive.

Figure 13 gives related results of full-wave computations of ionospheric penetration losses for plane waves, over the 15 kHz-500 KHz range, under nominal day- and night- ionospheric conditions. Northern hemisphere, mid-latitude parameters were assumed; and, the direction of wave propagation was parallel to the geomagnetic field (i.e., for the geomagnetic parameters chosen, the plane waves were propagating to the south, with an incidence angle of 20 degrees). Again (as shown later), this corresponds to optimum ionospheric penetration conditions.

Similar to the previous example, the normal daytime penetration losses



increase relatively sharply with increasing wave frequency. For the nighttime case, however, the losses are small, even for frequencies well above a few hundred kilohertz. This suggests, for example, that it should be possible under the appropriate nighttime conditions to receive Loran-C navigation signals (100 kHz) and standard AM broadcast signals (in the lower band around 600 kHz) in space.

### 3.2. Computations Relating to the Coupling Issue

As described earlier in conjunction with Figures 10 and 11, the large differences between the NAA penetration data acquired at Eglin and Ft. Churchill (there was little, if any, penetration at Ft. Churchill) prompted computational investigations of some of the directional issues associated with long wave coupling into the ionosphere. Figure 14 gives the results of calculations to investigate this issue further (magnetic dip = 70 degrees; wave incidence angle = 20 degrees). At the lower VLF frequencies, the calculations show only a few dB differences between waves propagating to the north and to the south; however, the computed differences become quite large as the wave frequency increases. For example, at 100 kHz for this nighttime example, the penetration loss is only a few dB for propagation to the south, but is more than 20 dB for propagation to the north. Hence, if one was to look for 100 kHz signals in space from a Loran C station, for example, it appears that they might be observable only southward of the transmitter. At higher frequencies, the differences in north/south penetration losses shown in Figure 14 are even more dramatic and severe.

### 3.3. Incidence Angle Effects

Some of the computed results

discussed above were for what was termed "best" or "optimum" penetration conditions. As illustrated by the results given in Figure 14, one optimum condition is wave propagation to the magnetic south for a northward directed geomagnetic field. Another optimum condition relates to the incidence angle of the waves on the ionosphere, as illustrated in Figure 15, for which penetration losses were computed for southward propagating waves as a function of incidence angle on the ionosphere. For the 15 kHz, 100 kHz, and 300 kHz examples shown, the penetration losses were least for an incidence angle near 20 degrees, which corresponds to a plane wave propagating parallel to the geomagnetic field (whose dip angle was assumed to be 70 degrees for the computations). In all cases, the penetration losses increase for incidence angles departing from that optimum condition. The penetration losses become progressively steeper and more prohibitive as the wave frequency increases and as the incidence angles vary away from the minimum-loss condition. The results indicate that although frequencies well above 100 kHz from ground-based transmitters may penetrate the nighttime ionosphere under suitable propagation conditions, the location(s) at which they might be received in space may very well be very highly restricted.

### 3.4. Tracking Long Waves Through the Ionosphere

The computational results described above represent numerical solutions for infinite plane waves incident on a horizontally uniform model ionosphere. However, the waves from a real long wave transmitter are divergent, and might be more accurately described as spherical rather than plane. There is then a question as to the realism of plane wave solutions in certain real

scenarios. The basic mathematics for investigating this question already exist (e.g., Ref 5), but the procedures required are not easy to implement. However, some progress has been made on a related question: "what path would a collimated beam of long waves follow in the ionosphere?"

As outlined schematically in Figure 16, a simple beam can be constructed by combining the fields of two plane waves of the same amplitude and frequency but with slightly different angles of incidence. For a given field component, say the transverse magnetic field, this gives an interference pattern in the vertical plane. By tracing the interference maxima a beam path or "wave track" is obtained. The process has a classical analog in the theory of Group Velocity in which two waves in the same direction but with slightly different frequencies are combined to give an interference pulse.

Figure 17 illustrates the results of "tracking" the maximum of the combined fields of two 15 kHz plane waves, one propagating with an incidence angle of 40 degrees, the other with an incidence angle of 50 degrees, as they penetrate a nighttime ionosphere. If the combined fields represent, effectively, the "track" of a plane wave incident on the ionosphere at an angle of 45 degrees, the computation indicates that the wave undergoes an appreciable off-set as it penetrates the ionosphere from the ground. It must be noted that the validity of this "wave-tracking" concept and computational approach has not been subjected to rigorous testing. Nevertheless, the "off-set" issue raised here is an intriguing one and it, along with some of the other results and issues raised in earlier Figures, indicate that further investigation is needed to fully characterize the long wave ionospheric transmission window.

#### 4. CONCLUDING COMMENTS

The discussions above have focussed on experimental and computational studies relating to the penetration of long radio waves up-through the ionosphere. Most of the results and issues raised, however, relate to the penetration of long radio waves down-through the ionosphere, as well. One primary issue for either case are the effective losses incurred in penetration of the ionosphere by long waves. These depend upon a number of factors as illustrated by the results shown earlier. Via reciprocity arguments, the effective losses should be the same for up- or down-going plane waves that are propagating parallel to the geomagnetic field. As indicated earlier, the losses can increase dramatically for non-parallel propagation, depending on wave frequency. The results indicate that ELF/VLF radio waves can penetrate the ionosphere with relatively little loss; and, under more restrictive conditions, LF, and even MF waves up to about 500-600 kHz, can penetrate the ionosphere without suffering prohibitive absorption.

Another key issue involves the coupling of signals out-of or in-to the earth-ionosphere waveguide. As illustrated earlier, the coupling out-of the waveguide in the northern hemisphere (where the geomagnetic field is pointing northward) is most efficient for waves propagating to the south. Alternately, reciprocity considerations indicate that the excitation of northward propagating waves in the earth-ionosphere waveguide by waves penetrating down-through the ionosphere may be more efficient than the excitation of southward propagating waves. The issue pertains to potential coverage-areas on or near the surface of the earth that would be provided by long wave sources in space.

Detailed knowledge of many long wave penetration issues is still lacking due, primarily, to the lack of a sufficiently large and diverse set of experimental data. Overall, there is a paucity of pertinent, quantitative, data from satellite monitoring of ground-based long wave transmissions; and, to date, no man-made long wave transmissions from space have been received on the ground. If plans for future experiments in these areas are carried out successfully, this will hopefully change so that the basic question of where in space do long wave transmissions from the earth reach (or, where on the earth will long wave transmissions from space be accessible) will be adequately addressed and answered.

#### REFERENCES

1. Leiphart, J.P., Zeek, R.W., Bearce, L.S., and E. Toth, "Penetration of the ionosphere by very-low-frequency radio signals- interim results of the LOFTI I experiment", Proc.IRE, 50(1), pp 6-17, 1962.
2. Bullough, K., Hall, A.J., Hall, J.E., and R.W. Smith, "Rocketborne equipment for an L.F. D-region propagation experiment", Proc.IEE, 114(6).703, 1967.
3. Inoue, Y., and S. Horowitz, "Numerical solution of full-wave equation with mode coupling", Radio Science, 1(8), pp 957-970, 1966.
4. Booker, H.G., Crain, C.M., and E.C. Field, "Transmission of Electromagnetic Waves Through Normal and Disturbed Ionospheres", RAND R-558-PR, Santa Monica, CA, Nov. 1970.
5. Budden, K.G., "Radio Waves in the Ionosphere", Cambridge, UK, Cambridge University Press, 1961.

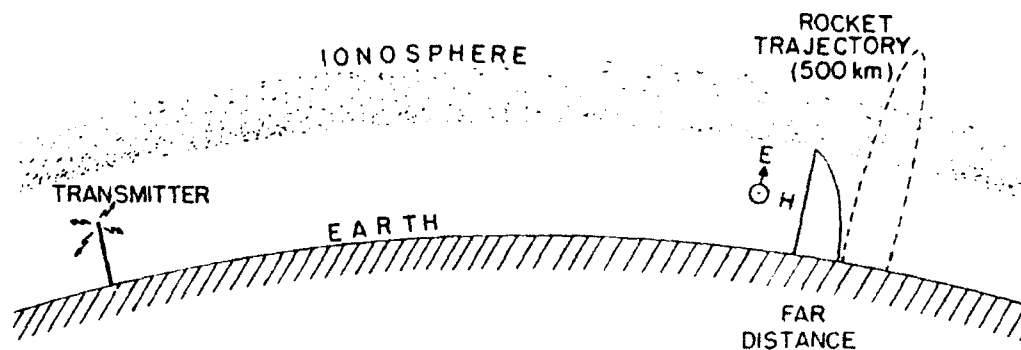


Figure 1. Rocket Probe Investigations of VLF Ionospheric Transmission

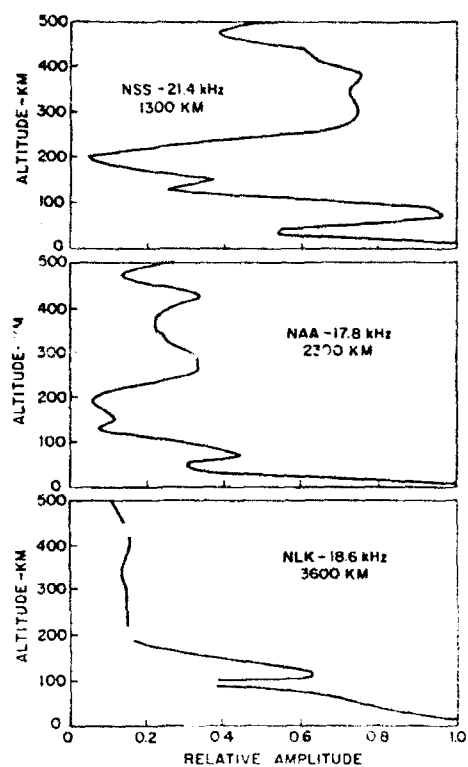


Figure 2. Eglin VLF Penetration Data:  
Night

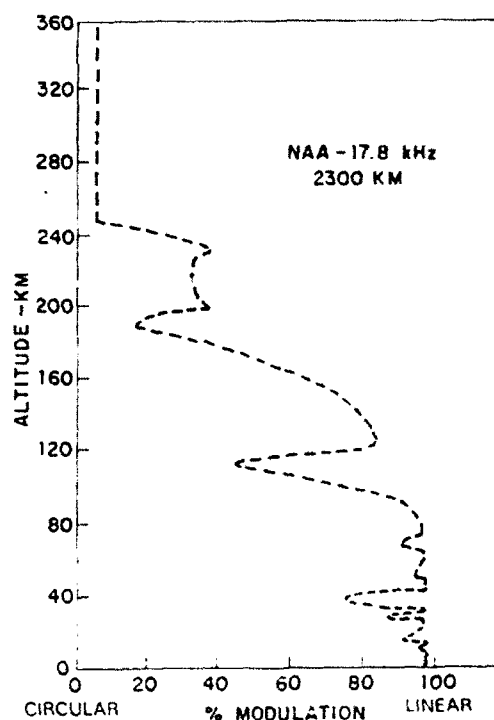


Figure 3. Polarization Profile:  
Eglin Night

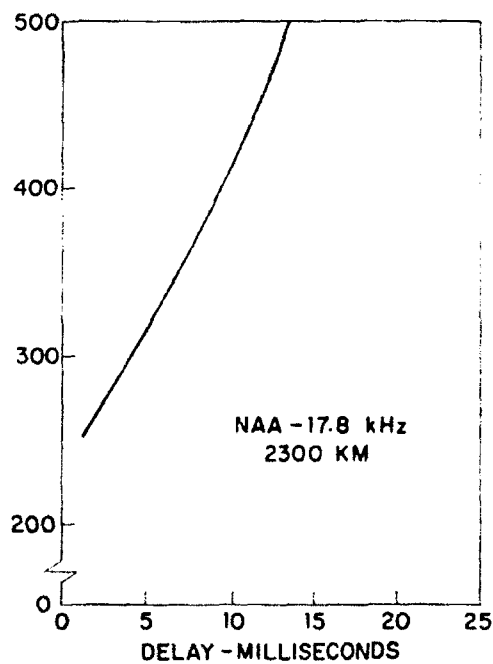
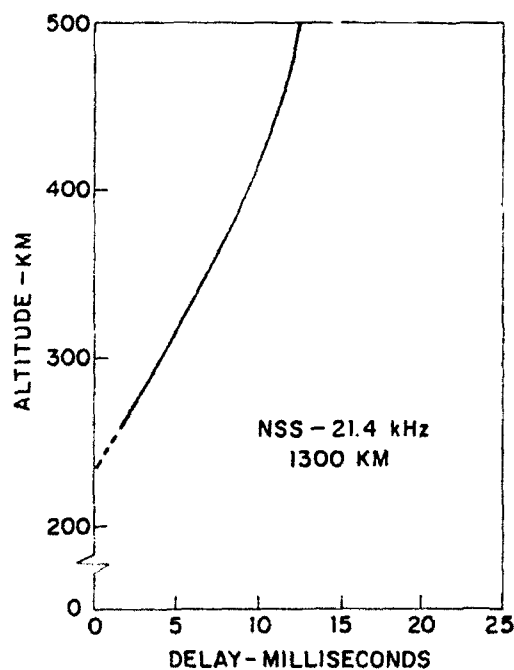


Figure 4. Signal Delays: Eglin Night

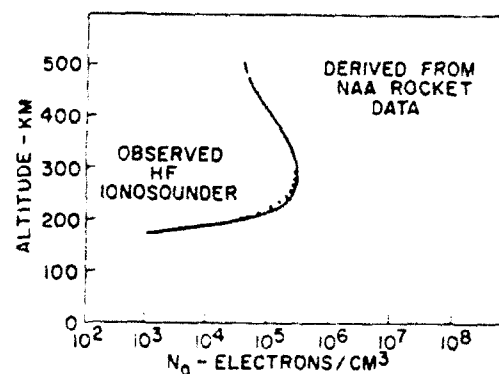
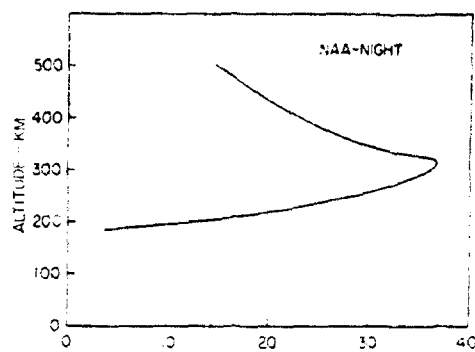
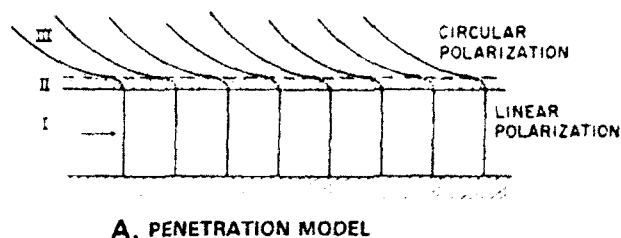
**B. DEDUCED REFRACTIVE INDEX PROFILE - EGLIN NIGHT****C. ELECTRON DENSITY PROFILES - EGLIN NIGHT**

Figure 5. Refractive Index/Electron Density Profiles  
Deduced from NAA Penetration Data

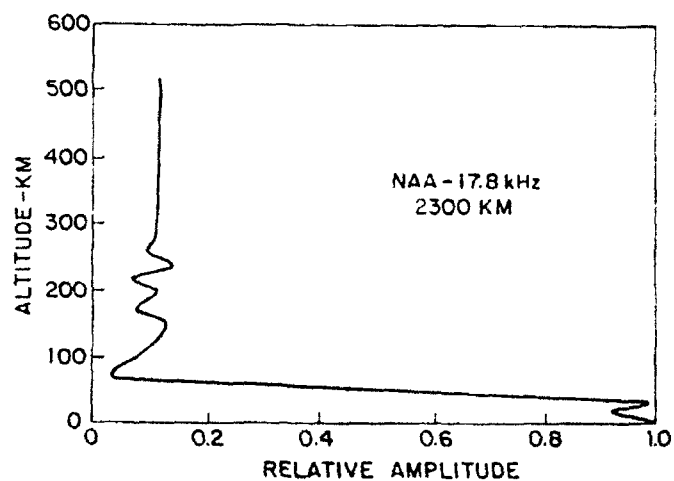
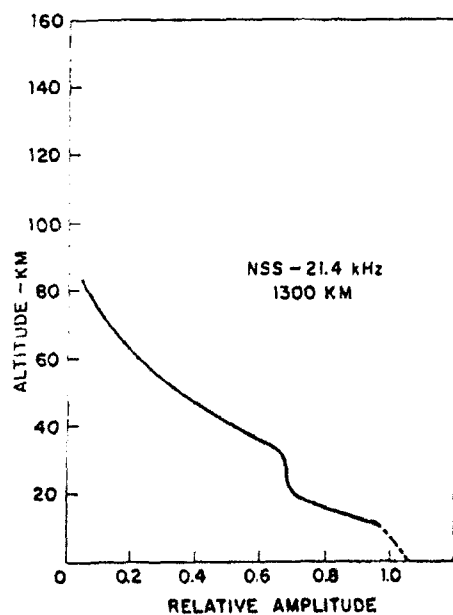


Figure 6. Eglin VLF Penetration Data: Day

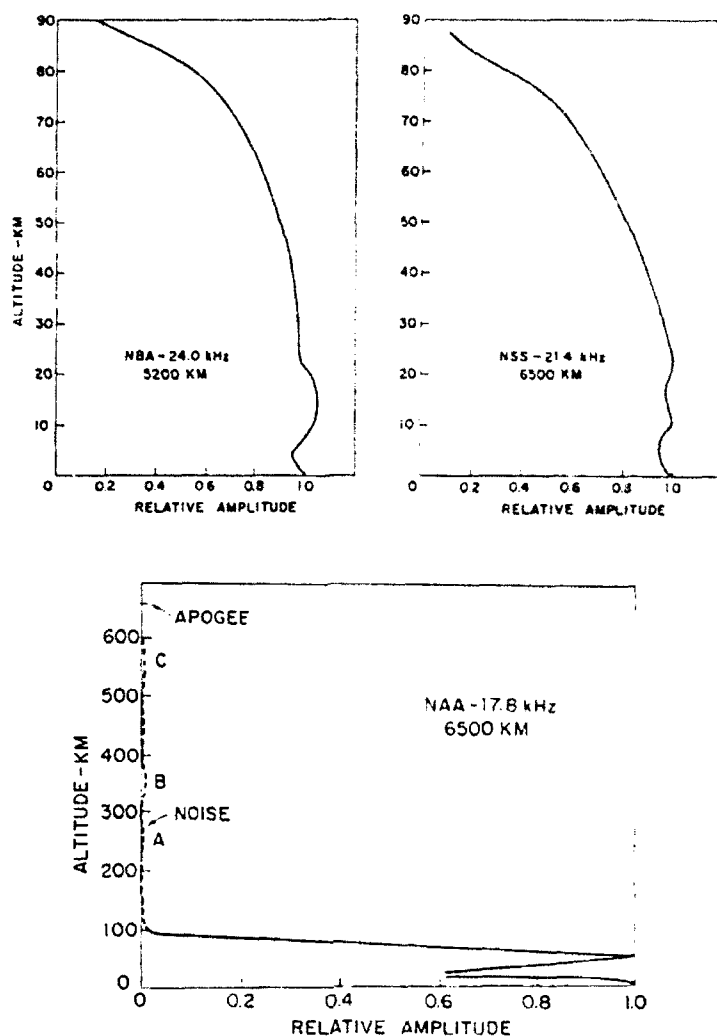


Figure 7. Natal VLF Penetration Data: Day

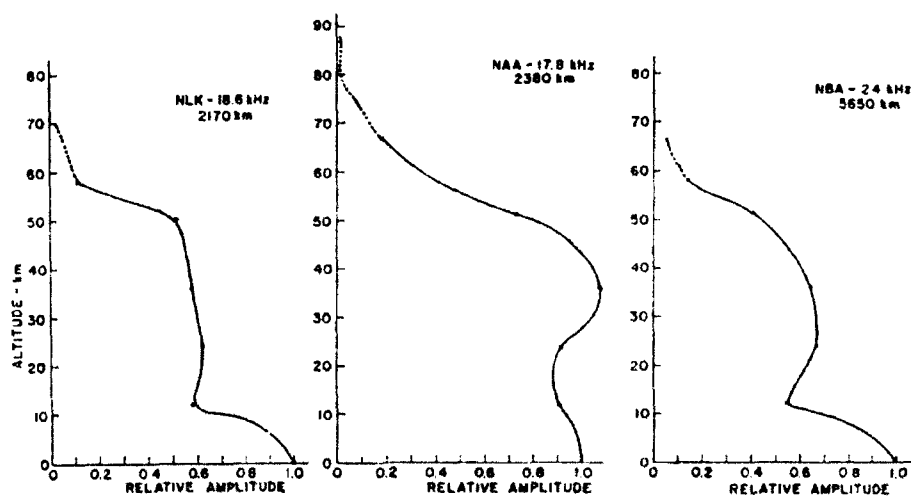


Figure 8. Ft. Churchill VLF Penetration Data: Day

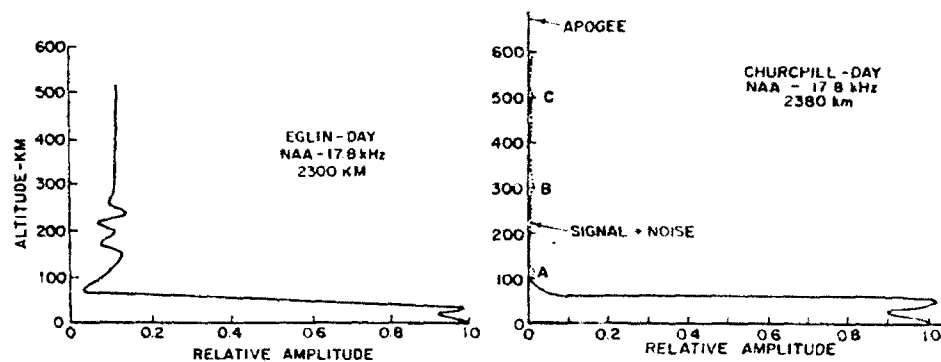


Figure 9. Comparison of Eglin and Ft. Churchill VLF Ionospheric Penetration Data

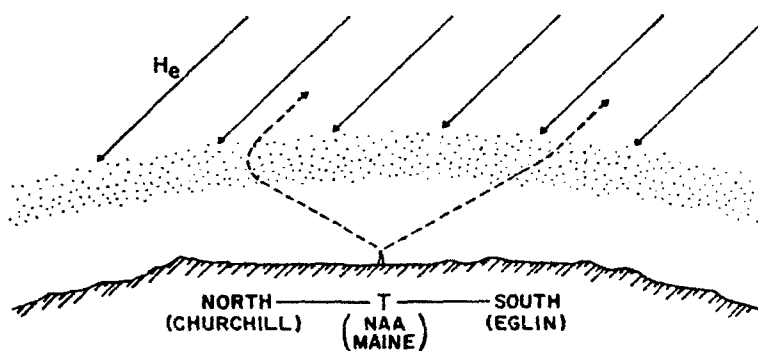


Figure 10. Coupling Issue

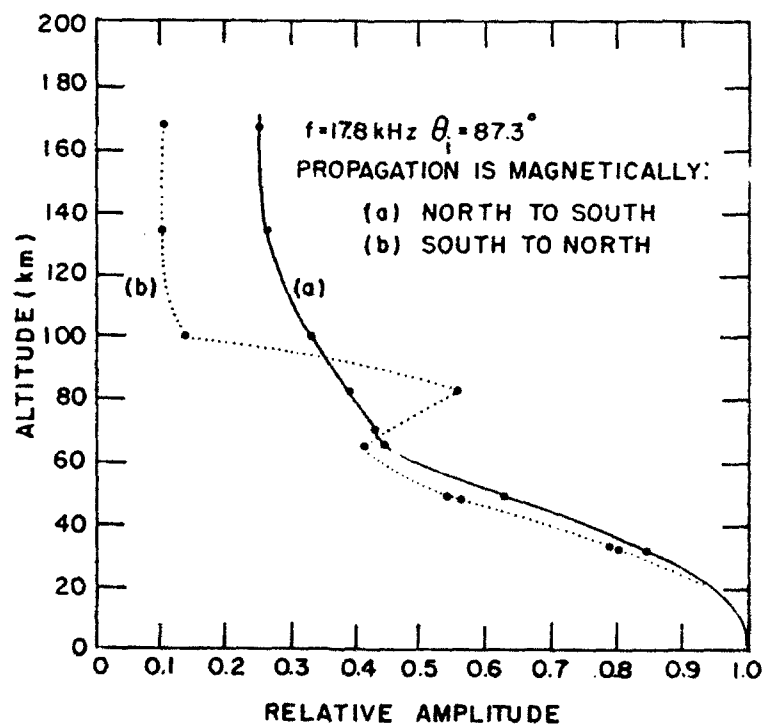


Figure 11. Calculated Example of a Directional Effect on VLF Ionospheric Penetration

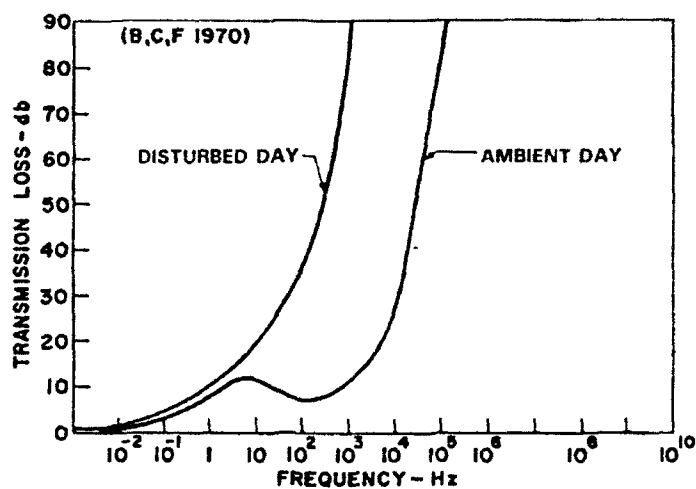


Figure 12. Long Wave Ionospheric Penetration Losses:  
Normal/Disturbed Conditions

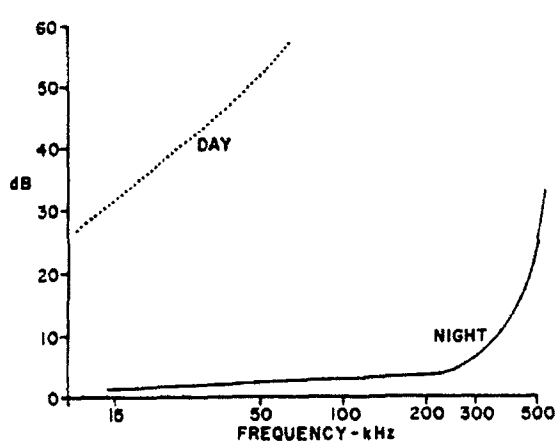


Figure 13. Long Wave Ionospheric  
Penetration Loss: Optimum  
Case

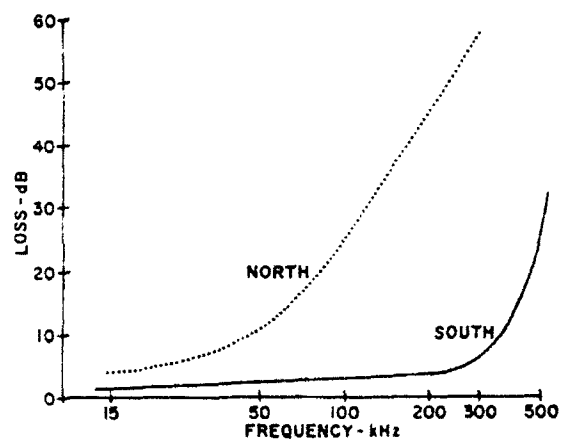


Figure 14. Long Wave Ionospheric  
Penetration: Directional  
Effects

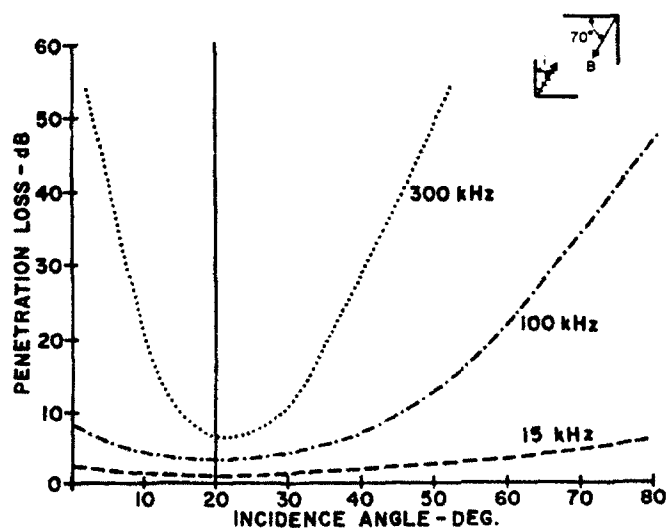


Figure 15. Long Wave Ionospheric Penetration:  
Incidence Angle Effects



- \* RAY OPTICS: NOT ADEQUATE FOR VLF
- \* GROUP THEORY: TWO WAVES-SAME INCIDENCE ANGLE/DIFFERENT FREQUENCIES
- \* WAVETRACK: TWO WAVES-SAME FREQUENCY/DIFFERENT INCIDENCE ANGLES

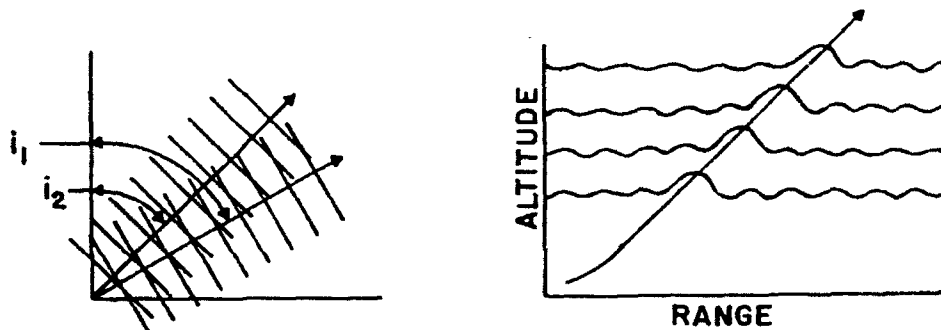


Figure 16. WAVETRACK Concept

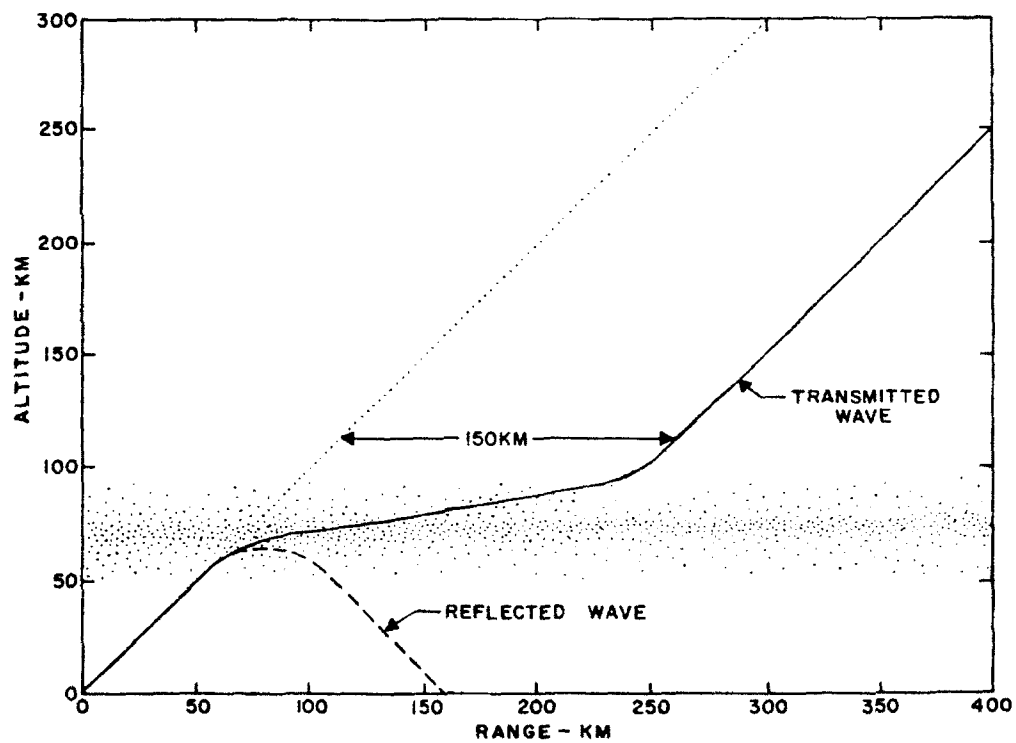


Figure 17. WAVETRACK Example

## DISCUSSION

### J. WAIT

1. Did you deal explicitly with the source in the full wave calculation of the penetration loss?
2. Also, I think it would be interesting to apply reciprocity to the case where the source is in the ionosphere and/or the magnetosphere.

### AUTHOR'S REPLY

1. No. We took an admittedly simplified approach to investigate relative effects relating to the penetration of long waves up through the ionosphere. Specifically, we computed the penetration of infinite plane waves, incident on a horizontally stratified ionosphere, for a wide range of propagation conditions. To deal appropriately with the source, a much more rigorous approach, such as outlined in works by Budden, Galejs, and Wait, for example, should be done. We were not computationally set up to handle such a treatment.
2. I agree; however, another somewhat thorny issue arises when the source is to be embedded in a magnetized plasma. Do we understand the interactions and coupling between it and the medium well enough to adequately describe its radiation characteristics (patterns, etc.)? This issue certainly was of great concern to U.S. Navy researchers when they were contemplating the development of a satellite-borne long-wave transmitter twenty years ago or so, and the issue still remains a difficult one to resolve today as well.

### M. PITTEWAY

At around, say, 10 kHz, we would expect a refractive index of 10 or 20 in the whistler medium. If we have an airborne VLF transmitter, therefore, transmission to the earth would only be possible through a narrow cone of angles. Other angles must be reflected or absorbed, so most of the transmitted energy will be lost to us? (The refraction gives an earth-based VLF transmitter an effective polar diagram of just a few degrees in the whistler medium. The airborne transmitter enjoys no such advantage).

### AUTHOR'S REPLY

The computations discussed in the talk centered on the upward penetration of infinite plane waves incident on a horizontally stratified ionosphere. Although the focus was on waves incident from below, reciprocity considerations show that the results can be used to obtain information relating to plane waves penetrating down through the ionosphere as well. Your comment actually deals with the same one Dr. Wait has just brought up: what signals will be received in space (or on the ground) for real long-wave sources on the ground (or, in space)? As mentioned earlier, those much more vigorous computations are beyond the scope and emphasis of the present paper.

# THE EFFECTS OF MULTIPATH SCATTER FROM WIND DRIVEN GRAVITY WAVES ON LINES OF BEARING SERVING AS DIRECTION FINDERS

by

R.H. C. It  
General Research Corporation  
1601 Randolph Road SE  
Suite 200 S  
Albuquerque, NM 87106  
United States

## ABSTRACT

Unambiguous locations for target transmitters are based on lines of bearing (LOB's) obtained by radio direction finding (DF) methods. These target location estimates based on intercepted LOB's are subject to error if multipath exists. This multipath may result from skywave interference or seascatter from rough seas. Skywave multipath becomes important for ranges greater than 300 km while ground wave multipath may be significant for ranges less than about 100 km. Although, the groundwave loss may be greater than the one-hop skywave mode, it may be comparable to the loss associated with the two-hop mode. Therefore both groundwave and skywave multipath may contribute to LOB errors for a given fleet exercise. For the case of skywave multipath, newer DF systems will be able to differentiate skywave signals from ground wave signals when the elevation angle is greater than about 18°.

In this paper groundwave multipath is estimated using the compensation theorem. This method provides a convenient analytical tool for relating the change in mutual impedance between a transmitting monopole and a receiving loop over a rough sea surface and a flat sea. The change in mutual impedance, in turn, is related to the change in current in each of the receiving loops used to generate LOB's.

Some predicted and observed examples of fleet exercise LOB errors from wind generated swell will be given based on target transmitter frequency, wind direction and speed of the ship carrying the DF system.

## 1 INTRODUCTION

The ability to detect low-flying missiles is of crucial concern to modern naval fleets. Unambiguous locations for target transmitters are based on lines of bearing (LOB's) obtained by radio direction finding (DF) methods. These target location estimates are subject to error if rough sea multipath exists. This paper addresses the errors in LOB's when sea multipath occurs. The mathematical approach uses the Monteath Compensation Theorem [1] relating the change in mutual impedance between the transmitting monopole and receiving loop antennas.

## 2 ANALYSIS

The total  $z$ -component for the Hertz vector for the geometry in Figure 1 is

$$\begin{aligned} \Pi_z(\omega) = \frac{I ds}{4\pi i \omega \epsilon_0} & \left\{ \underbrace{\frac{e^{ikr_1}}{r_1}}_{\text{direct wave}} + \underbrace{\left( \frac{\cos\theta - \delta}{\cos\theta + \delta} \right) \frac{e^{ikr_2}}{r_2}}_{\text{reflected wave}} \right\} \\ & \underbrace{\hspace{10em}}_{\text{"space wave"}} \\ & - \underbrace{\left( \frac{2\delta}{\cos\theta + \delta} \right) \frac{e^{ikr_2}}{r_2} \left( \frac{1}{2\Omega} + \frac{1 \cdot 3}{4\Omega^2} + \frac{1 \cdot 3 \cdot 5}{8\Omega^3} + \dots \right)}_{\text{"Norton Surface Wave"}} \end{aligned}$$

where

$$\delta = \begin{cases} \frac{1}{\sqrt{\epsilon}}, & \text{vertical polarization} \\ \sqrt{\epsilon}, & \text{horizontal polarization} \end{cases}$$

$$\epsilon = \epsilon_r + \frac{i\sigma}{\omega\epsilon_0} \quad (2)$$

$$\Omega = \frac{ikr_2}{2} \delta^2 \left( 1 + \frac{\cos\theta}{\delta} \right)^2 = \frac{ikr_2}{2} (\cos\theta + \delta)^2 \quad (3)$$

and the time dependence is  $e^{-i\omega t}$ . For the case of an antenna on an OUTBOARD ship, the angle  $\theta$  in (1) is near  $\frac{\pi}{2}$  and the direct and reflected waves cancel leaving only the Norton surface wave to propagate the signal from the target ship. The Norton surface wave in (1), normalized to the free-space field is

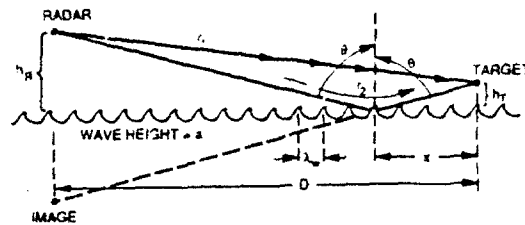


Figure 1. Geometry of Sea Multipath

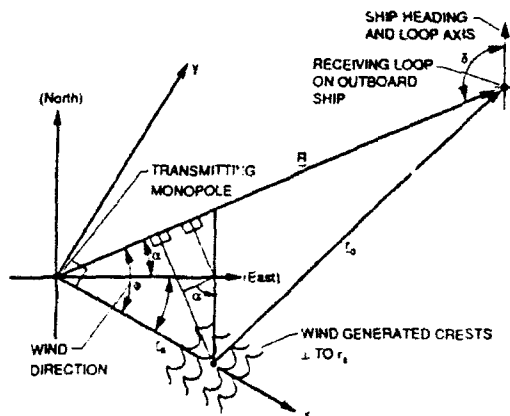


Figure 2. Geometry of transmitting monopole on target ship and receiving loop on outboard ship.

$$\text{Norton Surface Wave} \cong \frac{2}{ikD\delta^2} \quad (4)$$

where we have used only the leading term in equation (1).

From equation (4), the magnetic field of a transmitting monopole on the target ship is approximated by

$$H_x \cong \frac{h_A I_A \sin \alpha}{2\pi r_S^3} \left( \frac{2}{ik\delta^2} \right) (1 - ikr_S) e^{ikr_S} \quad (5)$$

$$H_y \cong \frac{h_A I_A \cos \alpha}{2\pi r_S^3} \left( \frac{2}{ik\delta^2} \right) (1 - ikr_S) e^{ikr_S} \quad (6)$$

which accounts for induction field effects of the monopole over the rough sea and

$$r_S = \sqrt{x^2 + y^2 + h_A^2} \quad (7)$$

where  $h_A$  is the monopole height above sea water.

Figure 2 shows the geometry of the transmitting monopole and the receiving loop on the OUTBOARD ship in frame  $x, y, z$ . The fields of the receiving loop are given approximately by (assuming the loops axis is nearly perpendicular to the  $z$ -axis; i.e., the loop is electrically close to the surface of the sea).

$$E_x \cong 0 \quad (8a)$$

$$E_y \cong 0 \quad (8b)$$

$$E_z \cong \frac{ik^2 \eta_o I_B S}{2\pi R} \sin \delta e^{ikr_o} \quad (8c)$$

where

$$r_o \cong R - x \cos \phi \quad (8d)$$

$$\phi = \alpha + \text{wind-direction} \quad (8e)$$

and  $\eta_o = 120 \pi$  ohms,  $I_B$  is the current in the loop,  $S$  is the loop's effective area, and  $\delta$  is the angle the loops horizontal axis makes with the vector  $R$  in Figure 2.

Using the following first-order approximation to the complete compensation theorem; i.e., assuming  $E$  is the unperturbed field and  $H$  is the perturbed field [2], yields the change in mutual impedance between the loop and the monopole.

$$\Delta Z = Z'_{AB} - Z_{AB} = \frac{1}{I_A I_B} \int_S (E \times H) \cdot e_n da \quad (9)$$

where  $E$  is the field of the receiving loop,  $H$  is the field of transmitting monopole, and  $E$  and  $H$  are defined above in (5), (6) and (8), and  $Z_{AB}$  is the mutual impedance between the monopole and loop when the sea is smooth; i.e.,

$$Z_{AB} = -\frac{\eta_o k^2 S h_A \sin \delta}{2\pi R} e^{ikR}, \quad (kR \gg 1) \quad (10)$$

and the unit normal  $e_n$  to the rough sea is

$$e_n = \frac{-e_x \zeta'(x) + e_y}{[1 + (\zeta')^2]^{1/2}} \quad (11)$$

with  $\zeta'(x)$  the local slope of the rough sea surface.

Substituting (5), (6), (8) and (11) into (9) gives

$$\frac{\Delta Z}{Z_{AB}} = (ik \cos \phi / 2) \int_{x_0}^{x_1} \zeta'(x) \exp(-ikx \cos \phi) H_1^{(1)}(kx \cos \phi) dx, \quad (12)$$

The limits of integration,  $x_0$  and  $x_1$ , are such that we can use the asymptotic expansion for the Hankel function,  $H_1^{(1)}(z)$ , as

$$H_1^{(1)}(kx \cos \phi) \sim \left( \frac{2}{\pi kx \cos \phi} \right)^{1/2} e^{-i3\pi/4} e^{ikx \cos \phi} \quad (13)$$

and substituting in (12) yields the result

$$\frac{\Delta Z_{AB}}{Z_{AB}} = \begin{cases} e^{-i\pi/4}, & |\phi| \leq \pi/2 \\ e^{i\pi/4}, & |\phi| > \pi/2 \end{cases} \int_{x_0}^{x_1} \frac{\zeta'(x)}{\sqrt{x}} dx, \quad (14)$$

Wait [3], has obtained a similar result to (14).

### 3 EQUIVALENT CIRCUIT FOR DETERMINING LOOP CURRENTS

The two-loop equivalent circuit representing the transmitting monopole and receiving loop make the following assumptions in the circuit parameters.

$$R_{rad}^{mono} \simeq \ll R_{rad}^{mono} \text{ (electrically short monopole)} \quad (15a)$$

$$R_{rad}^{mono} \simeq 400 \left( \frac{h_A}{\lambda} \right)^2 \quad (15b)$$

$$\omega L_{mono} = \frac{60[\ln(h_A/\lambda) - 1]}{\tan(kh)} \quad (15c)$$

$$R_{rad}^{loop} = 20 (k^2 \pi a^2)^2 \simeq 10^{-2} \Omega,$$

$$\text{for } a/\lambda = 1/60$$

$$[4, \text{ p. 182, fig. 5.5}] \quad (15d)$$

$$R_{loss}^{loop} = R_{rad}^{loop} \quad (15e)$$

$$L_{loop} = a\mu_0 \left[ \ln \left( \frac{16a}{d} \right) - 2 \right]$$

$$a = \text{loop radius} \\ d = \text{wire diameter}$$

$$[4, \text{ pp. 183, 197-198}], [6] \quad (15f)$$

$$\begin{aligned} \omega M &= Z_{AB} \\ &= - \frac{\eta_0 k^2 S h_A \sin \phi}{2\pi R} e^{ikR} \end{aligned} \quad (15g)$$

In (15b) and (15c),  $h_A$  is the height of the transmitting antenna above the ground plane. In (15c), 60 rather than 120 occurs in the numerator because the image of the vertical antenna yields a factor of two. In the equivalent circuit, a capacitor is used in series with the terminating loop resistor to cancel the effect of  $L_{loop}$  and make the antenna resonant, [4], pp. 183, 197-198]. Using the results in (15), the equivalent circuit becomes a coupled RL circuit with loop equations given by

$$\begin{aligned} R_R^M I_A - i\omega L_m I_A + i\omega |M| I_B &= V \\ i\omega |M| I_A + (2R_R^L + R_T) I_B &= 0 \end{aligned} \quad (16)$$

with a determinant  $\Delta$ ,

$$\begin{aligned} \Delta &= \begin{vmatrix} (R_R^M - i\omega L_m) & i\omega |M| \\ i\omega |M| & (2R_R^L + R_T) \end{vmatrix} \\ &= (R_R^M - i\omega L_m) (2R_R^L + R_T) + \omega^2 |M|^2. \end{aligned} \quad (17)$$

Solving for  $I_B$ , the current in the receiving loop, gives

$$I_B = \frac{-i\omega |M| V}{(R_R^M - i\omega L_M)(2R_R^L + R_T^M) + \omega^2 |M|^2}$$

or

$$I_B = \frac{-iZ_{AB}V}{(R_R^M - i\omega L_M)(2R_R^L + R_T^M) + |Z_{AB}|^2} \quad (18)$$

and

$$\frac{\Delta I_B}{I_B} = \frac{\Delta Z}{Z_{AB}} \quad (19)$$

The change in mutual impedance is therefore directly related to the change in current in the receiving loop on the OUTBOARD ship. These currents are used to generate LOB's, from "look-up tables".

From Figure 2

$$e_R = R / |R| \quad (20)$$

$$\cos \delta = e_R \cdot e_b \quad (21)$$

where  $e_b$  is a unit vector along the ship's center line and

$$\cos \phi = e_x \cdot e_R \quad (22)$$

### 3.1 Comparison of Predicted and Observed Bearing Errors

Measured LOB's were obtained from the fleet exercise experiment during the period 27 January to 6 February, 1990, in the Caribbean, typically from 2300 hours to 0800 hours local time. There were 43 targeting events for the Ship A and 43 for the Ship B using Ship C a cooperative mobile target (CMT) at ranges from 20 to 100 nmi.

From equations (14) and (19), the error in the current  $I_2$ , in each of the receiving loops on the OUTBOARD ship is a function of  $\zeta(x)$  which in turn is a function of the sea state or wind speed. These waves are referred to as swell when they continue on after the wind stops or changes direction. The wave characteristics are very nearly

in the shape of an inverted cycloid, with the vertical distance between trough and crest the wave height, and the horizontal distance between successive crests called the water wave length,  $\lambda_w$ . Figure 3 shows a statistic related to the wave height, i.e.,  $H_{1/3}$ , or the so-called significant wave height for four different spectrums for the wave surface [7]. Reference [8] gives photos showing the ocean waves developed when the wind has been blowing a sufficient time (1 knot = 0.514442 m/sec.).

Using the Gauss-Legendre numerical algorithm, the integral in Equation (14) and, hence the phase of the current in the receiving loop, was evaluated for the parameter values shown in Table I.

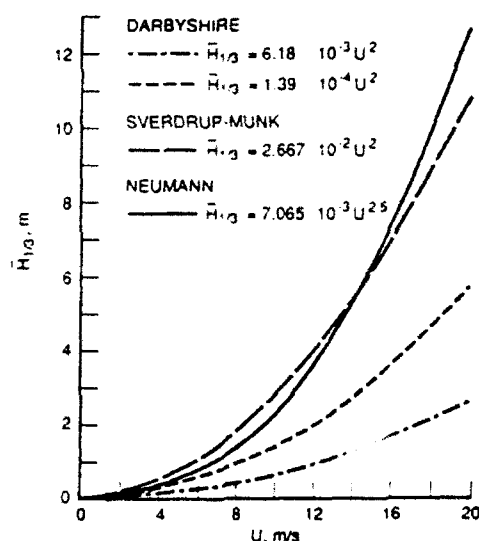


Figure 3. Significant Wave Height,  $H_{1/3}$ , versus Wind Speed using 4 different Spectra ([7]).

Table I. Input parameters for predicted LOB errors for Ship A, Event 2.

Azimuth (from CMT to OUTBOARD ship relative to North, deg.)	-44°
Distance between CMT and OUTBOARD ship (nmi)	29
Wind direction (relative to east, deg.)	50°
Wind velocity (knots)	12
Frequency (MHz)	20.18
Predicted LOB error (deg.)	1.3°

Figure 4 shows the geometry for Event 2 and the LOB multipath correction for rough sea multipath. Another method for correcting for rough sea multipath would be during the calibration testing phase where appropriate modifications can be made to the look-up table (database) for the magnitude and phase of the currents in the loop antennas on the OUTBOARD ship. However, Figure 4 suggests the possibility of correcting for rough sea multipath in "real time" without modifying the look-up table.

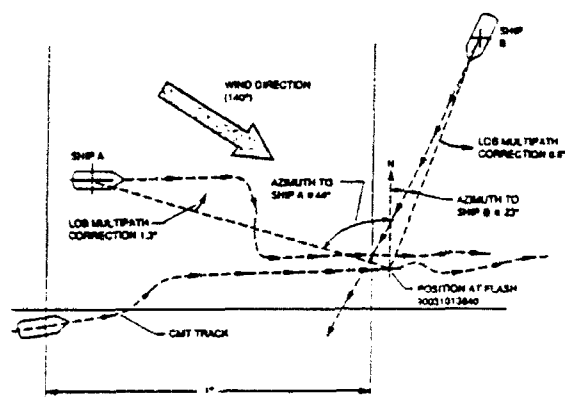


Figure 4. Geometry for Event 2 and LOB multipath correction.

Figure 5 shows a plot of the difference of the two LOBs in Figure 4 versus time. Except for the time period 2.55 to 2.60, the LOB error was positive; i.e., the LOB from the correlation Interferometer Direction Finding CIDF algorithm was greater than the LOB predicted using the dead-reckoning algorithm. The dead-reckoning algorithm, developed by this author, predicts the accurate "ground-truth" LOB using the time "flashes" at approximately 10 minute intervals and an interpolation scheme for intermediate times, to determine the distance the target ship traveled along a great circle path from its initial position at latitude,  $\theta_i$ , longitude,  $\phi_i$ , and time flash  $t_p$  to a new latitude and longitude. Using the parameters for the sea state shown in Table I, a predicted LOB multipath error of about 1.3° was obtained. Actually, the predicted LOB multipath error in Figure 5 is very close to the average LOB error (i.e., 0.93°). The average LOB error is only significant in analyzing the accuracy of the LOB correction and not as a basis for LOB error prediction. The reason is the multipath error from rough seas is not constant with time, but changes as the sea state varies and until the wind has persisted for a sufficient duration to create swell. As it turned out, for Figure 5, the wind did remain

in a nearly constant direction at a relatively constant speed (12 knots) during the course of LOB readings. The large changes in the LOB error from time 2.55 to 2.65 may have been caused by factors other than rough sea multipath; e.g., and incorrect LOB from a target ship.

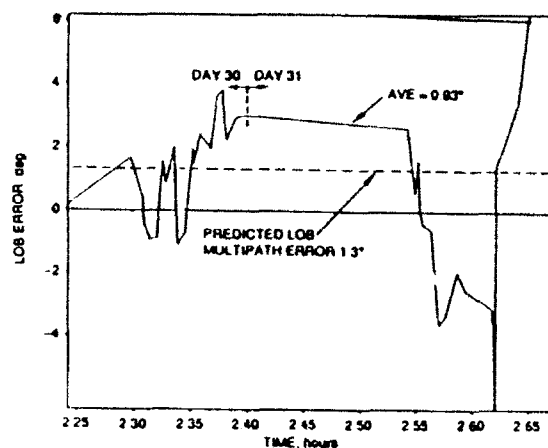


Figure 5. Event 2, Ship A, LOB difference error versus time. Also shown is the predicted LOB multipath error.

#### 4 CONCLUSIONS

A general expression for computing the change in loop current in the receiving loop antennas used in the OUTBOARD DF system is given as a function of sea state. This expression was evaluated for the sea's which existed during a fleet exercise and gave good agreement for the predicted and observed LOB errors.

## 5 REFERENCES

- [1] G.D. Monteath, "Application of the compensation theorem to radiation and propagation problems," *Proc. Inst. Elec. Eng.*, vol. 98, pp. 23-30, 1951.
- [2] R.H. Ott and J.R. Wait, "First order effects of terrain on the radiation pattern of a nondirectional LF beacon," *AEU*, vol. 27, pp. 106-110, eq. 11, 1973.
- [3] J.R. Wait, *EM Wave Theory*. New York: Harper and Row/Wiley, 1986, ch. 7.
- [4] C.A. Balanis, *Antenna Theory, Analysis and Design*. New York: Harper and Row, 1982, pp. 58, 182, fig. 5.5, pp. 196, 183, 197-198.
- [5] E.C. Jordan, *Electromagnetic Waves and Radiating Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1950, p. 312, eq. 10-54, p. 512, eq. 14-4.
- [6] W.L. Weeks, *Antenna Engineering*. New York: Mc-Graw-Hill, 1968, p. 60.
- [7] B. Kinsman, *Wind Waves, Their generation and propagation on the ocean surface*. Dover, New York, 1984, p. 391.
- [8] Beaufort Wind Scales, in the Appendix of Bowditch, Nathaniel, "American Practical Navigator, An Epitome of Navigation;" U.S. Navy Hydrographic Office under the authority of the Secretary of the Navy, U.S. Government Printing Office, Washington, D.C., pp. 727-735, 1962.

## DISCUSSION

### J. BELROSE

The experiment you described employed vertical monopoles for transmitting and small ferromagnetic loops for receiving on large metal ships/large with respect to the frequency used. The analysis you presented was based on well understood propagation parameters (the total signal being the vector sum of a direct and a ground reflected wave and a Norton surface wave). You addressed the coupling between the vertical transmitting antenna and the receiving loop antenna(s), coupled through radiation. But the radiation properties of these antennas are influenced by the conducting surface of the ships. The lower half of the monopole is the entire ship; the small receiving loop will couple to currents flowing on the conducting hull of the ship. Did you not experience problems from these effects?

### AUTHOR'S REPLY

I agree with the discussor's comment that indeed the entire ship is the radiator; i.e., the ship's superstructure modifies the currents excited in each of the loop antennas. However, this effect is included in the CIDF algorithm (correlation interferometer direction finding) when measured loop currents are correlated against the data base currents; the latter are obtained as the ship steams around a buoy and collects target data from a known target direction.



## PROPAGATION EFFECTS ON HF DIRECTION FINDING

George H. Millman  
Research Associates of Syracuse, Inc.  
510 Stewart Drive  
N. Syracuse, NY 13212 USA

## SUMMARY

Estimates are made of the propagation errors that can be made in the measurement of the geographic coordinates of an HF emitter when utilizing HF direction finding techniques. The sources of the errors considered in this analysis are those due to (1) imprecise knowledge of the ionospheric reflection height, (2) ionospheric tilts, i.e., gradients of electron density and (3) tropospheric refractive bending.

## 1.0 INTRODUCTION

The determination of the geographic coordinates of an HF radio/radar transmitter by direction finding is influenced to a great extent by the characteristics of the propagation media, i.e., the ionosphere and the troposphere.

Various techniques are available for locating an HF emitter utilizing two or more receiving sites. A single receiving site can also be employed for HF direction finding. The single site configuration requires that both the azimuth and elevation angle of the received signals be accurately measured. For the single site, the ionospheric height of reflection of the HF transmitted signals is required for the calculation of the ground distance to the emitter.

Errors in locating the coordinates of an HF emitter arise, in the case of the ionosphere, from (1) imprecise knowledge of the ionospheric height of reflection, (2) the presence of an electron density gradient, i.e., ionospheric tilt along the path of propagation and (3) the presence, along the propagation path, of a traveling ionospheric disturbance (TID), i.e., large scale electron density perturbation in the F region. An additional source of error results from the refractive bending phenomenon encountered in the troposphere. The presence of an ionospheric tilt and TID imparts errors in the measurement of both the emitter's bearing and ground range. The effect of ionospheric reflection height inaccuracy and tropospheric refraction is to introduce an error solely in the ground range measurement.

In this paper, estimates are made of the magnitude of the angular deviations and ground range errors that could be imposed on an HF direction finding system by the ionosphere and the troposphere. The

angular deviations and ground range errors are presented as a function of solar-geophysical and meteorological conditions.

## 2.0 THEORETICAL CONSIDERATIONS

## 2.1 Reflection Height Inaccuracy

The lack of precise data on the electron density variation with altitude at the ionospheric reflection point of an HF transmitted radio wave is the basis for considering the reflection height inaccuracy as a source of error in HF direction finding. An error in the height of reflection imposes an error in the estimation of the surface range to an HF transmitter.

The ground distance error,  $\Delta S$ , depicted in Figure 1, is simply

$$\Delta S = S - S_A \quad (1)$$

where  $S$  is the true ground distance. The parameter,  $S_A$ , is the apparent ground distance which is expressed by

$$S_A = 2r_0 \cos^{-1} \{ [1/2r_0(r_0 + h_A)] \} \quad (2)$$

$$[r_0^2 + (r_0 + h_A)^2 - R_{1A}^2]^{1/2}$$

where  $r_0$  is the radius of the earth,  $h_A$  is the apparent reflection height and  $R_{1A}$  is the apparent slant distance to the reflection point. It is assumed that  $R_{1A} = R_1$  where  $R_1$ , the true slant distance to the reflection point, is given by

$$R_1 = \{r_0^2 + (r_0 + h)^2 - 2r_0(r_0 + h) \cos(S/2r_0)\}^{1/2} \quad (3)$$

and where  $h$  is the correct reflection height.

## 2.2 Ionospheric Tilt

Electron density gradients present along the path of HF transmissions can cause the signals to undergo an angular deviation off the great circle path. In addition, a gradient can introduce a modification in the ground scatter distance.

To facilitate the analysis, it is assumed the propagation signal undergoes a plane mirror-type reflection in the ionosphere. The ionospheric tilt geometry employed in

this analysis is depicted in Figures 2 and 3. The tilt angle,  $\phi$ , is the angle between the vertical to the earth's surface and the normal to the tilted surface while the horizontal rotation angle,  $\theta$ , is the angle in the horizontal plane through which the tilted angle is rotated away from the great circle plane.

For the ionospheric tilt geometry shown in Figure 2, the deviation in azimuth,  $\Delta A$ , brought about by a tilt can be expressed by the function

$$\Delta A = \sin^{-1} \{ \sin(S_T/r_o) / \sin(S_{TL}/r_o) \} \quad (4)$$

where  $S_{TL}$  is the ground distance traversed by the deviated signals and  $S_T$  is the transverse component of  $S_{TL}$ .

It can be readily shown from spherical geometry that

$$S_{TL} = r_o \cos^{-1} \{ \cos(S_L/r_o) / \cos(S_T/r_o) \} \quad (5)$$

where  $S_L$ , the component of  $S_{TL}$  along the great circle path, i.e., the longitudinal component, can be expressed by

$$S_L = (S/2) + r_o \{ (\pi/2) - (E_L + i + 2\phi_i) \} \quad (6)$$

The parameter  $S$  is the ground scatter distance that would be attained in a nontilted ionosphere,  $E_L$  is the elevation angle in the longitudinal plane of the ray incident on the earth,  $i$  is the angle of incidence of the ray with respect to the vertical and  $\phi_i$  is the longitudinal component of the tilt angle,  $\phi$ . It is noted that the term,  $(i + 2\phi_i)$ , is the angle of reflection in the longitudinal plane.

The angle  $E_L$  can be determined from

$$E_L = \cos^{-1} \{ [(r_o + h)/r_o] \sin(i + 2\phi_i) \} \quad (7)$$

and the angle  $i$  from

$$i = \sin^{-1} \{ (r_o/R_u) \sin(S/2r_o) \} \quad (8)$$

where  $h$  is the reflection altitude and  $R_u$  is the slant range to the reflection point and is defined by

$$R_u = \{ r_o^2 + (r_o + h)^2 - 2r_o(r_o + h) \cos(S/2r_o) \}^{1/2} \quad (9)$$

Assuming that the tilted reflecting surface is spherical, it follows that the angle  $\phi_i$  can be expressed by

$$\phi_i = \sin^{-1} \{ \sin \phi \cos \theta \} \quad (10)$$

Referring to Equation (4), the transverse component of the ground scatter distance,  $S_T$ , is given by

$$S_T = r_o \{ (\pi/2) - (E_T + 2\phi_T) \} \quad (11)$$

where  $E_T$  is the transverse elevation angle of the reflected ray incident on the earth and  $\phi_T$  is transverse component of the tilt angle.

It can be shown that

$$E_T = \cos^{-1} \{ [(r_o + h)/r_o] \sin 2\phi_T \} \quad (12)$$

and

$$\phi_T = \sin^{-1} \{ \sin \phi \sin \theta \} \quad (13)$$

It is of interest to note the similarity between Equations (6) and (11), (7) and (12), and (10) and (13).

The error in the ground scatter distance,  $\Delta S_{TL}$ , can be derived from the expression

$$\Delta S_{TL} = S_A - S_{TL} \quad (14)$$

where  $S_A$  is the apparent ground scatter distance defined by the expression

$$S_A = 2r_o \cos^{-1} \{ [r_o^2 + (r_o + h)^2 - \{ (R_D + R_u)/2 \}^2 / 2r_o(r_o + h)] \}^{1/2} \quad (15)$$

The parameter,  $R_D$ , is the slant ray path distance between the ionospheric reflection point and the ground reflection point, and is expressed by

$$R_D = \{ r_o^2 + (r_o + h)^2 - 2r_o(r_o + h) \cos(S'/2r_o) \}^{1/2} \quad (16)$$

where  $S'$  is the surface distance between the subionospheric reflection point and ground reflection point. From spherical geometry, it can be shown that

$$S' = r_o \cos^{-1} \{ \cos \{ [S_L - (S/2)]/r_o \} \cos(S_T/r_o) \} \quad (17)$$

### 2.3 Tropospheric Refractive Bending

When electromagnetic waves are propagated through the troposphere and ionosphere, they undergo a change in direction or refractive bending. This phenomenon arises from the nonhomogeneous characteristics of the media. The tropospheric refractive bending when taken into account in HF propagation analysis is found to modify the ground scatter distance. This topic was initially discussed in a previous AGARD symposium (Reference 1) but is briefly reviewed in this paper for completeness of the subject pertaining to propagation effects on HF direction finding.

The effects of the troposphere on HF propagation can be evaluated by means of ray tracings. This requires that the index of refraction in both the troposphere and ionosphere be expressed as a function of altitude.

The index of refraction in the troposphere,  $n_t$ , is defined by

$$N = (n_t - 1) \times 10^6 = \frac{a}{T} \left( p + \frac{b\epsilon}{T} \right) \quad (18)$$

where  $N$  is the refractivity,  $T$  is the air temperature ( $^{\circ}\text{K}$ ),  $p$  is the total pressure (mbar) and  $\epsilon$  is the partial pressure of water vapor (mbar). According to Smith and Weintraub (Reference 2), the constants,  $a$  and  $b$ , are  $77.6^{\circ}\text{K/mbar}$  and  $4810^{\circ}\text{K}$ , respectively.

The tropospheric refractive index model employed in this analysis is the CRPL Reference Refractivity Atmosphere - 1958 (Reference 3) which is described by

$$N(h) = N_0 + (h - h_0) \Delta N \quad (19)$$

where  $N_0$  is the surface refractivity and  $h_0$  is the surface height above mean sea level. This expression is valid for  $h_0 \leq h \leq (h_0 + 1)$  km. The parameter,  $\Delta N$ , is defined by

$$\Delta N = -7.32 \exp(0.005577 N_0) \quad (20)$$

For the region contained within  $(h_0 + 1) \leq h \leq 9$  km, the refractivity decays as

$$N(h) = N_1 \exp[-c(h - h_1 - 1)] \quad (21)$$

where  $N_1$  is the value of  $N(h)$  at 1 km above the surface and

$$C = [1/(8 - h_1)] \log_e(N_1/105) \quad (22)$$

Above 9 km, the experimental decay is of the form

$$N(h) = 105 \exp[-0.1424(h - 9)] \quad (23)$$

The index of refraction in the ionosphere,  $n_i$ , can be expressed by the relationship

$$n_i = \left\{ 1 - \left( \omega_p / \omega \right)^2 \right\}^{1/2} = \left\{ 1 - \left( N_e e^2 / \epsilon_0 m_e \omega^2 \right) \right\}^{1/2} \quad (24)$$

where  $\omega_p$  is the angular plasma frequency of the medium (rad/s),  $N_e$  is the electron density (electrons/ $\text{m}^3$ ),  $e$  is the electron charge ( $1.6 \times 10^{-19}$  c),  $m_e$  is the electron mass ( $9.1 \times 10^{-31}$  kg),  $\epsilon_0$  is the electric permittivity of free space ( $10^{-9}/36\pi$  F/m) and  $\omega$  is the angular frequency of the incident wave (rad/s).

The distribution of electron density with height is assumed to follow the Chapman model of the form (Reference 4).

$$N_e = N_m \exp \left\{ (1/2) \left[ 1 - \left\{ (h - h_m) / H_e \right\}^2 \right] \right\} \exp \left\{ - \left[ (h - h_m) / H_e \right] \right\} \quad (25)$$

where  $N_m$  is the electron density at the level of maximum ionization,  $h_m$ , and  $H_e$  is the scale height of the neutral particles; i.e., the height of the homogeneous atmosphere at a given temperature.

It should be noted that the ionospheric refractive index, given by Equation (24), is also a function of both the electron collision frequency and the earth's magnetic field. According to Davies (Reference 4), the effect of the collision frequency term on the index of refraction is negligible for frequencies on the order of 10 MHz and above and at altitudes greater than 80 km. When neglecting the earth's magnetic field, an error is introduced in the computation of the index of refraction. The maximum error evaluates to less than 0.4 percent at a frequency of 30 MHz and slightly greater than 1 percent at 20 MHz (Reference 1). These calculations are based on the daytime electron density models presented in Table 1 and on a magnetic induction of 0.5 G which is assumed to be invariant with altitude (Reference 5).

### 3.0 DISCUSSION

#### 3.1 Reflection Height Inaccuracy

An error in the estimation of the ionospheric height of reflection of HF signals could arise from inaccurate data on the distribution of electron density at the midpoint of the transmission path. A source of the error could be a travelling ionospheric disturbance inducing a perturbation in the ionization.

An estimate of the ground distance error resulting from an error in ionospheric reflection height is shown in Figure 4. It is evident that an underestimation of the reflection height results in an overestimation of the ground distance. The effects are reversed when there is an overestimation of the reflection height. For a given reflection height error, the error in the ground distance decreases with increasing ground distance.

According to Figure 4, for an assumed reflection height error of -20 km and ground distance of 3400 km, the error in the prediction of the ground distance evaluates to approximately 12 and 13 km for reflection heights of 300 and 350 km, respectively. For a -10 km reflection height error and assuming identical conditions, the ground distance error is on the order of one-half the -20 km results.

#### 3.2 Ionospheric Tilt

The azimuthal deviations due to ionospheric tilts oriented at a horizontal rotation angle,  $\theta$ , of  $45^{\circ}$  (and  $135^{\circ}$ ) and  $90^{\circ}$  are presented in Figures 5 and 6, respectively, as a function of tilt angles,  $\phi$ , of  $1^{\circ}$ ,  $3^{\circ}$  and  $5^{\circ}$  and reflection heights of 300 and 350 km. The

lack of data beyond the ground distances of 2600 and 3000 km is attributed to the reflected rays not reaching the earth's surface, i.e., overshooting the earth.

An examination of the data discloses that the angular deviation decreases with increasing ground distance and increases with increasing tilt angle and reflection height. For a fixed set of conditions, i.e., tilt angle, reflection height and ground distance, the angular deviation maximizes at  $\theta = 90^\circ$ .

It should be noted that, for  $\theta = 0^\circ$  and  $180^\circ$ , i.e., the tilt oriented in the great circle plane, HF transmissions would not experience an angular deviation from the great circle plane.

The data shown in Figures 5 and 6 can be readily converted to spatial dimensions. According to Figures 7 and 8, there is a slight increase in the lateral displacement with ground distance. In the ground distance interval of 700 and 3400 km, the lateral displacement, for  $\theta = 5^\circ$  and  $\theta = 45^\circ$  and  $135^\circ$ , varies between approximately 37 and 38 km for  $h = 300$  km and between 44 and 45 km for  $h = 350$  km. When the horizontal rotation angle is  $90^\circ$ , the magnitude of the lateral displacement increases to approximately 53 and 56 km for a reflection height of 300 km and to approximately 62 and 65 km for  $h = 350$  km.

The ground distance errors resulting from the presence of ionospheric tilts are depicted in Figures 9 through 11 for horizontal rotation angles of  $0^\circ$  and  $180^\circ$ ,  $45^\circ$  and  $135^\circ$ , and  $90^\circ$ , respectively. It is apparent that for a fixed set of conditions, i.e., tilt angle, reflection height and ground distance, the ground distance error is a maximum at a horizontal rotation angle of  $0^\circ$  and  $180^\circ$  (Figure 9) and a minimum at  $90^\circ$  (Figure 11). For a tilt angle of  $5^\circ$ , reflection height of 350 km and ground distance of 2600 km, the ground distance error is approximately 17.6 km for  $\theta = 0^\circ$  and  $180^\circ$ , 10.5 km for  $\theta = 45^\circ$  and  $135^\circ$  and 0.7 km for  $\theta = 90^\circ$ .

It is of interest to note that for both horizontal rotation angles of  $0^\circ$  and  $180^\circ$  (Figure 9) and  $45^\circ$  and  $135^\circ$  (Figure 10) the ground distance error increases with increasing ground distance while, for a horizontal rotation angle of  $90^\circ$ , the reverse occurs; that is, the ground distance error decreases with increasing ground distance.

The reversal effect of the ground distance error is shown in Figure 12 which is a plot of the error as a function of various horizontal rotation angles for a tilt angle of  $5^\circ$  and reflection height of 350 km. It is evident that, as the angle  $\theta$  increases, the position of the minimum ground distance error is displaced towards increasing ground distance. For  $\theta = 0^\circ$ ,  $45^\circ$ ,  $60^\circ$  and  $90^\circ$ , the minimum error is located at a ground distance of

approximately 700, 900, 1100 and 3400 km, respectively.

### 3.3 Tropospheric Refractive Bending

The effects of tropospheric refraction on HF propagation is evident in the data tabulated in Table 2 which is a compilation of the computations presented at the 1979 Lisbon AGARD symposium (Reference 1). The Table 2 calculations are based on the four daytime electron density models described in Table 1 and on surface refractivities of 320 - and 400-N units. The electron density models and surface refractivities are inputs to the mathematical relationships defining the ionospheric and tropospheric index of refraction profiles, respectively, which are required for the ray tracing process.

In the analysis, it was assumed that (1) the troposphere is contained between the earth's surface and 30-km altitude, (2) the base of the ionosphere is located at an altitude of 80 km and (3) free space prevails in the region between the troposphere and ionosphere.

An examination of Table 2 reveals that, when the refractive characteristics of the troposphere are taken into account, HF radio waves, in general are propagated to a greater distance than when the troposphere is neglected as indicated by a surface refractivity of 0-N units. In addition, the true height of reflection decreases while the virtual height of reflection increases. However, in the case of transmissions at 20 MHz and  $1^\circ$  elevation angle, the reverse occurs; that is, the ground distance and virtual reflection height both decrease with increasing surface refractivity.

It should be noted that, when the ionospheric index of refraction is defined in terms of magnetic field parameters, the ground distance and the true and virtual reflection heights are found to be a function of both magnetic field conditions and surface refractivity (Reference 6).

An interesting disclosure of Table 2 is the existence of long range propagation paths in a portion of the 20-MHz data.

### 4.0 CONCLUSIONS

For an HF direction finding system, imprecise information on the height of reflection of ionospheric propagation HF radio waves produces an error in the measurement of the ground distance to an emitter. The error increases with both increasing reflection height and reflection height error and decreases with increasing ground distance. The presence of an ionization tilt in the region of ionospheric reflection can cause HF signals to undergo both an azimuthal bending and a change in the transmission path length which, in turn, would produce an error in the ground distance measurement.

The magnitude of the azimuthal deviation and the ground distance error is a function of the tilt angle, the tilt altitude, the ground distance and the orientation of the tilt with respect to the great circle plane of the HF transmissions.

The angular bending is zero for a tilt oriented in the great circle plane and a maximum when it is orthogonal to the plane. The angular deviation increases with increasing tilt angle and tilt altitude and decrease with increasing ground distance.

The lateral displacement which is the angular deviation expressed in terms of a spatial dimension increases slightly with increasing ground distance.

For a tilt oriented in the plane of the great circle path, the ground distance error is a maximum and increases with increasing ground distance. However, for the case in which the tilt is perpendicular to the plane, the error is a minimum and decreases with increasing ground distance. The ground distance error brought about by a ionospheric tilt is less than that resulting from an inaccurate estimation of the height of ionospheric reflection.

The phenomenon of tropospheric refractive bending tends to have the following effects on HF propagation: to decrease the true height of reflection and generally to increase the virtual height of reflection and to increase the ground distance over which the waves can be transmitted. However, geophysical conditions could exist which could result in the virtual reflection height and the ground distance decreasing with increasing severity of tropospheric refraction.

## 5.0 REFERENCES

1. Millman, G.H., "Tropospheric Effects on HF Propagation", Special Topics in HF Propagation, AGARD Conference Proceedings No. 263, pp (9-1) - (9-14), November 1979.
2. Smith, E.K. and S. Weintraub, "The Constants in the Equation for Atmospheric Refractive Index at Radio Frequencies", Proceedings of IRE, Vol. 41, pp. 1035 - 1037, August 1953.
3. Bean, B.R. and E.J. Dutton, "Radio Meteorology", National Bureau of Standards Monograph 92, US Government Printing Office, Washington, DC, 1966.
4. Davies, K. "Ionospheric Radio", IEE Electromagnetic Waves Series 31, Peter Peregrinus Ltd, United Kingdom, 1990.
5. Millman, G.H., "The Effect of Tropospheric Refraction on HF Propagation", General Electric Technical Information Series Report No. R77EMH5, March 1977.
6. Millman, G.H., "Refraction Effects on Magnetic Field Geometry and HF Propagation", Journal of Atmospheric and Terrestrial Physics, Vol. 37, pp. 751-760, May 1975.

Table 1

### Ionospheric Electron Density Profiles

Model	Layer	Scale Height Hs (km)	Altitude (km)	Maximum Electron Density $N_m$ ( $\times 10^{11}$ Electrons/m <sup>3</sup> )	Plasma Frequency (MHz)
A	E	10	100	1.500	3.477
	F <sub>1</sub>	40	200	3.000	4.917
	F <sub>2</sub>	50	300	12.500	10.037
B	E	10	100	1.459	3.429
	F <sub>1</sub>	40	200	2.918	4.850
	F <sub>2</sub>	50	300	12.160	9.899
C	E	10	100	1.440	3.406
	F <sub>1</sub>	40	200	2.879	4.817
	F <sub>2</sub>	50	300	11.997	9.833
d	E	10	100	1.421	3.384
	F <sub>1</sub>	40	200	2.841	4.785
	F <sub>2</sub>	50	300	11.838	9.767

Table 2  
Tropospheric and Ionospheric Ray Tracing Calculations

Frequency (MHz)	Elevation Angle (deg)	Surface Refractivity (N-units)	Average Ground Distance (km)	Average True Reflection Height (km)	Average Virtual Reflection Height (km)
10	1.0	0	1863.9	81.8	85.4
		320	1949.9	81.7	92.7
		400	1984.3	81.6	95.8
	3.5	0	1468.7	82.3	88.3
		320	1511.0	82.2	92.2
		400	1526.3	82.1	93.6
20	1.0	0	4820.3	196.5	546.8
		320	4729.0	186.4	522.5
		400	4282.6	167.7	457.6
	3.5	0	4231.4	215.9	514.8
		320	4438.8	215.5	564.4
		400	4521.6	215.4	585.5
30	1.0	0	3979.8	248.7	361.9
		320	4067.6	248.2	377.9
		400	4123.6	248.1	388.4
	3.5	0	3501.0	252.0	364.4
		320	3540.1	251.3	371.7
		400	3574.5	251.2	378.2

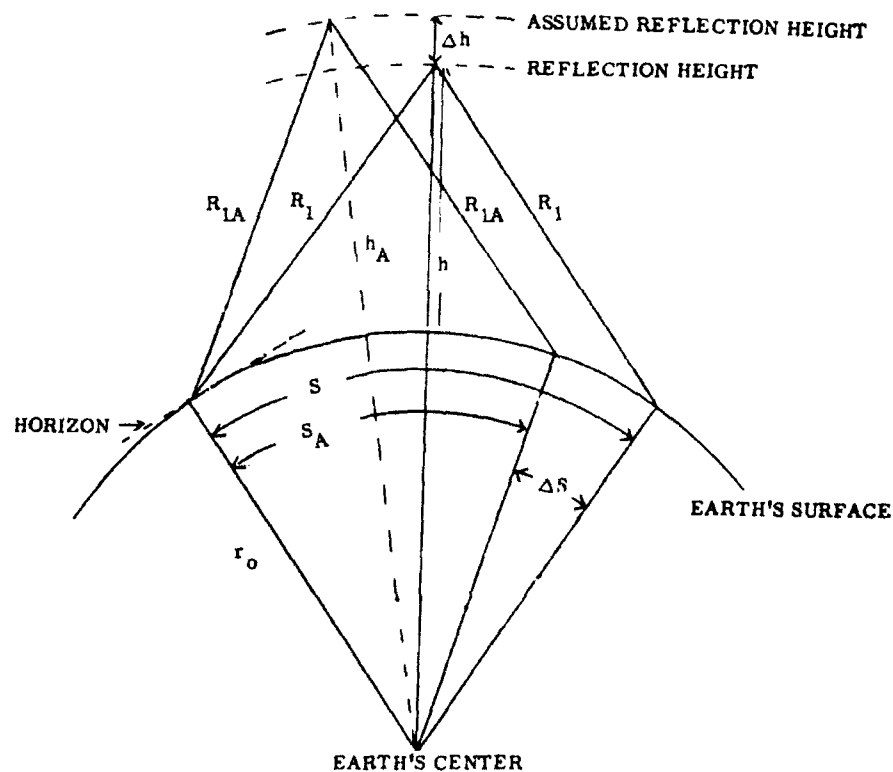


Figure 1. Geometry for Reflection Height Inaccuracy

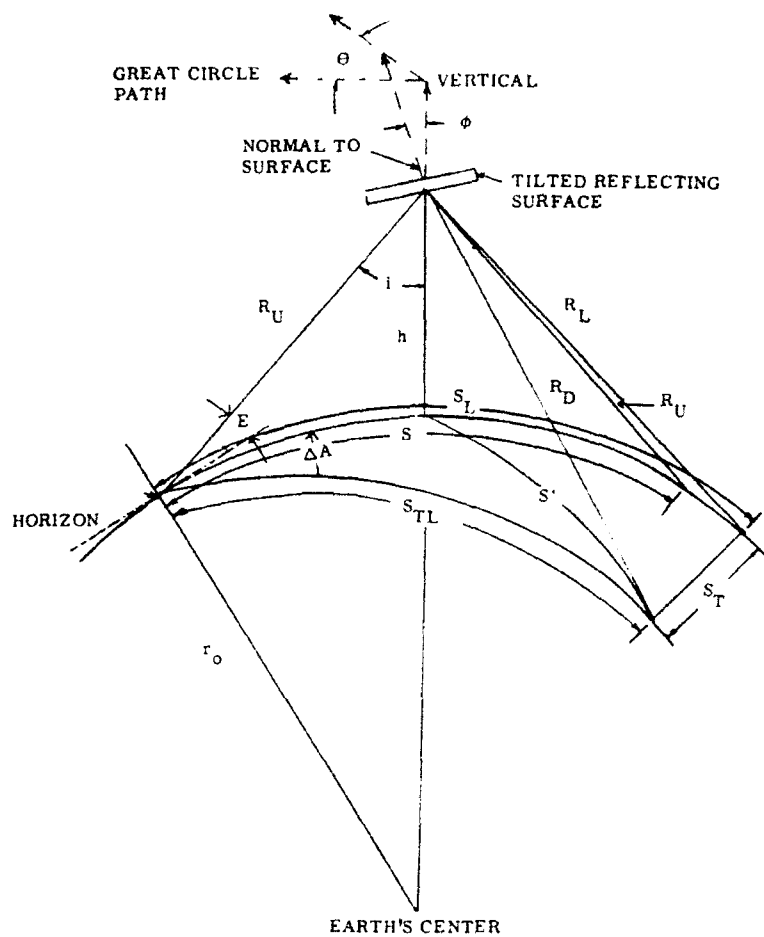
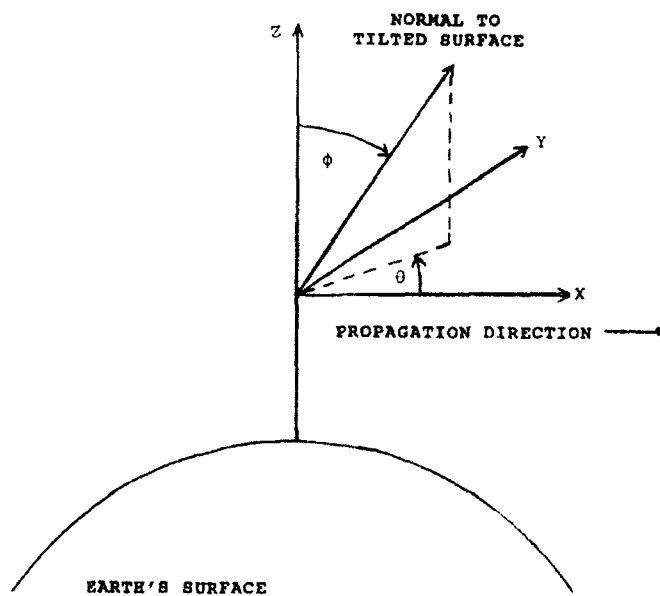


Figure 2. Ionospheric Tilt Geometry

Figure 3. Ionospheric Tilt Angle,  $\phi$ , and Horizontal Rotation Angle,  $\theta$

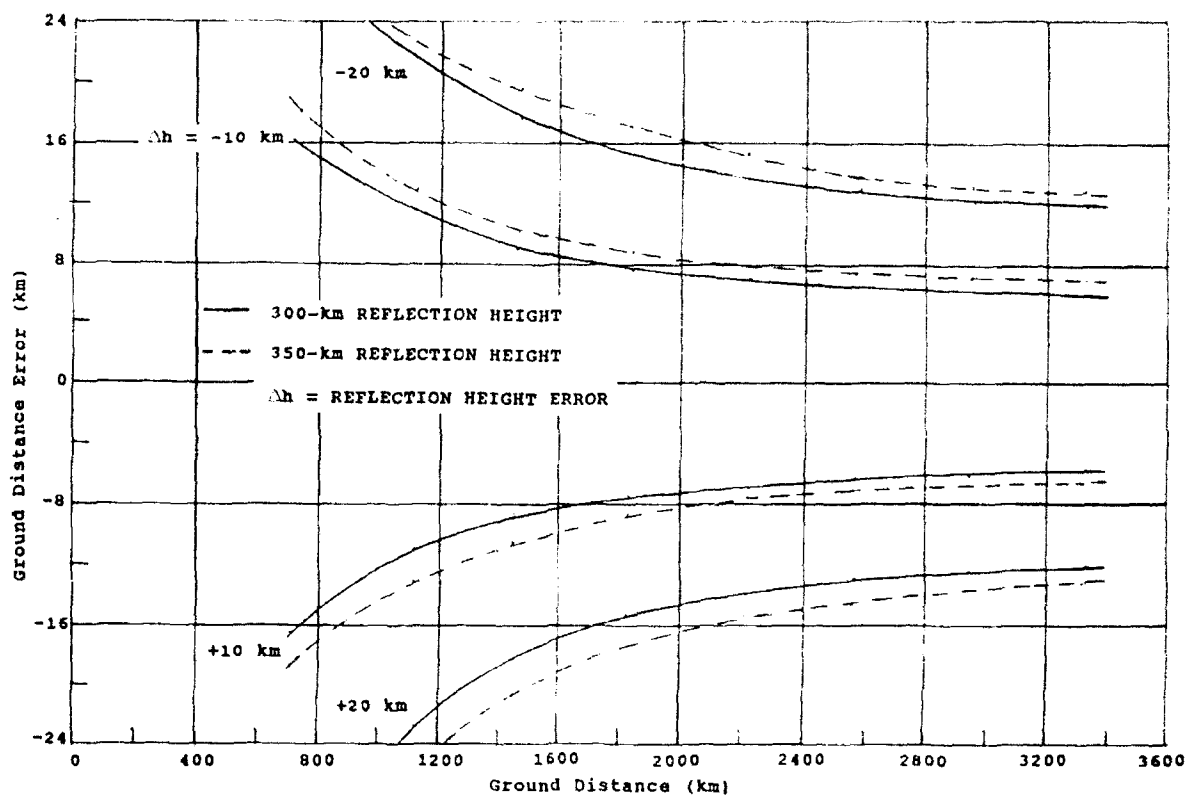
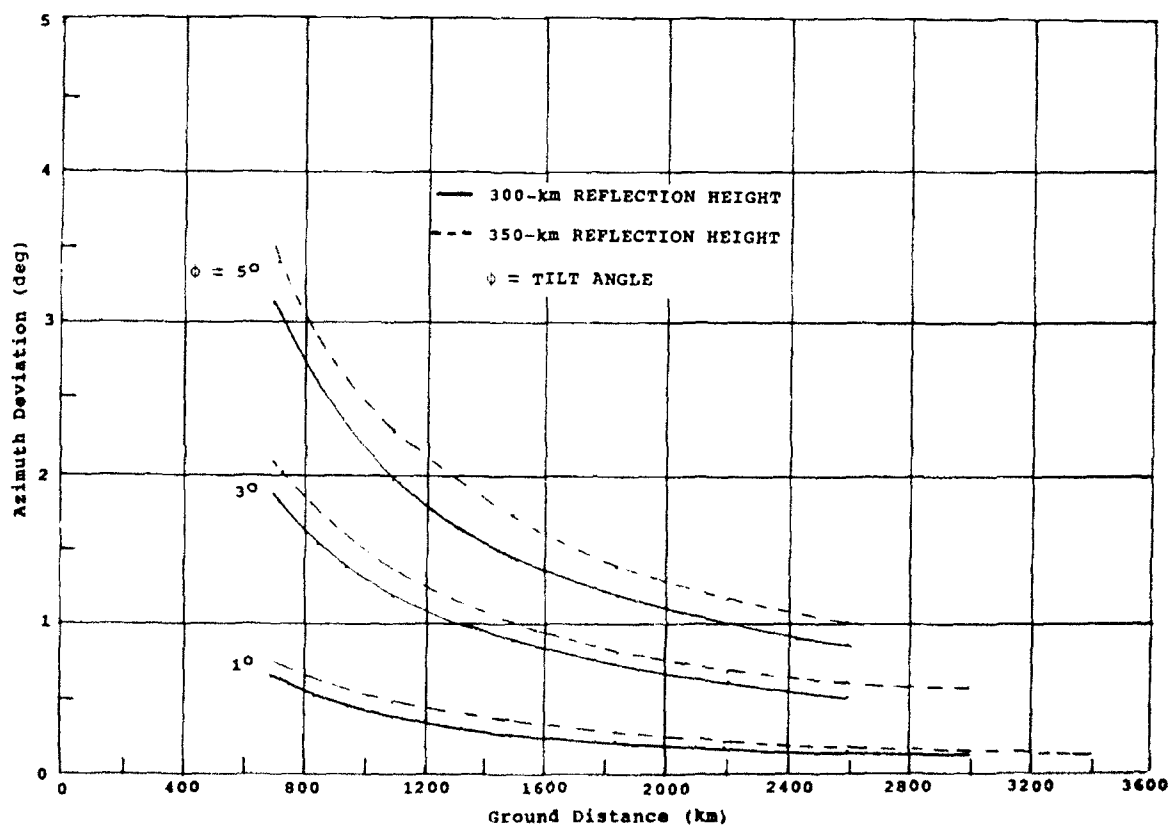


Figure 4. Ground Distance Error Due to Error in Reflection Height.

Figure 5. Azimuthal Deviation Due to Tilt at Horizontal Rotation of  $45^\circ$  and  $135^\circ$



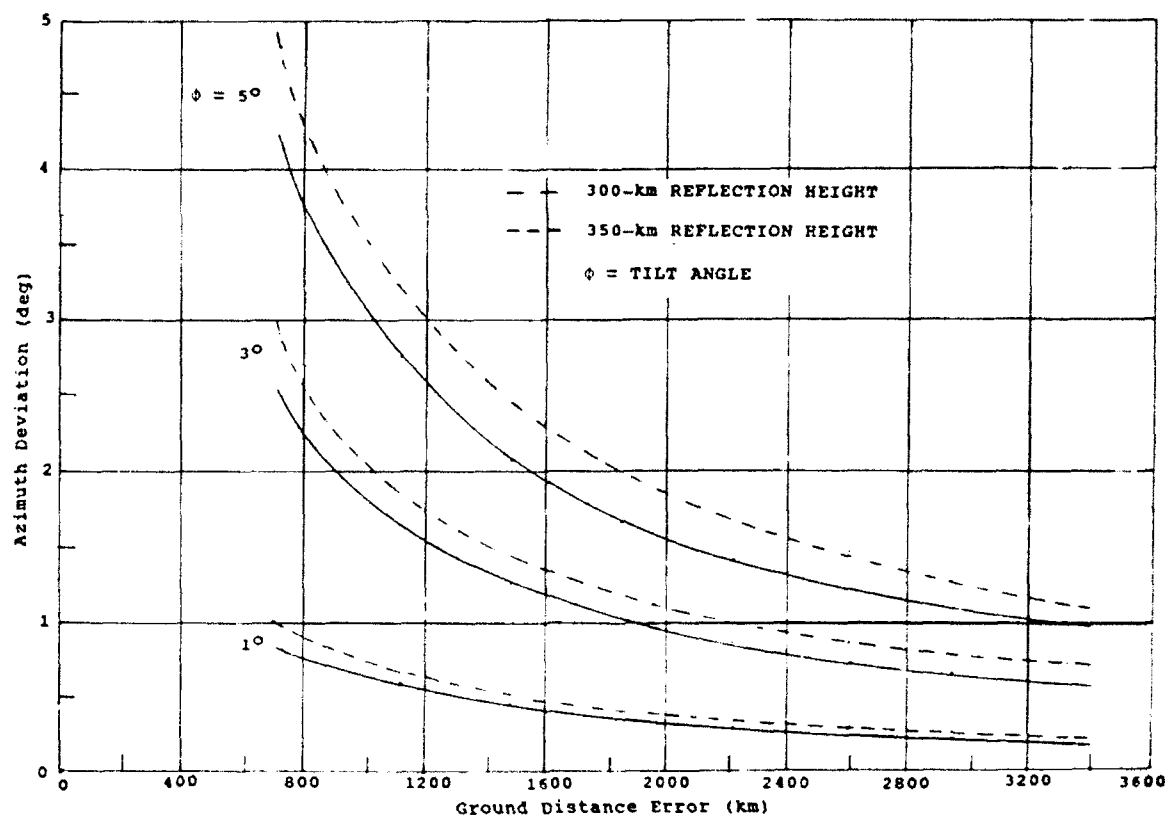


Figure 6. Azimuthal Deviation Due to Tilt at Horizontal Rotation of  $90^\circ$

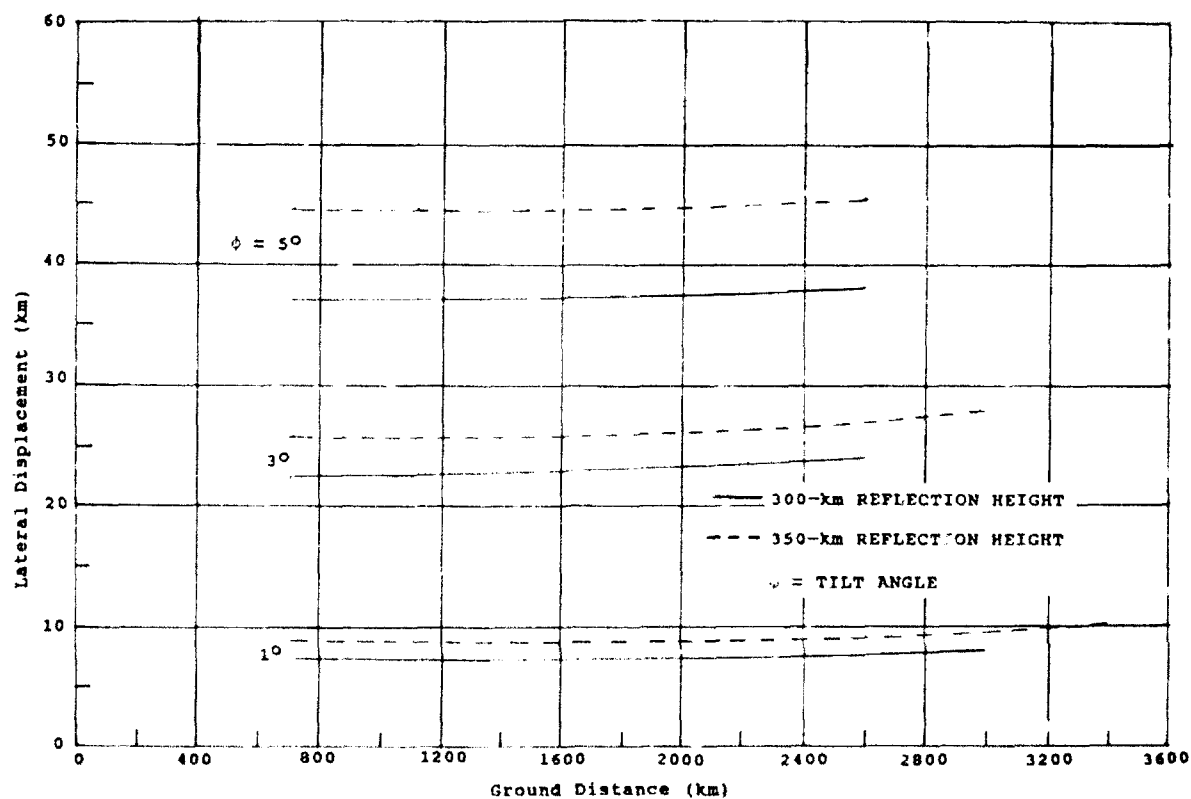


Figure 7. Lateral Displacement Due to Tilt at Horizontal Rotation of  $45^\circ$  and  $135^\circ$

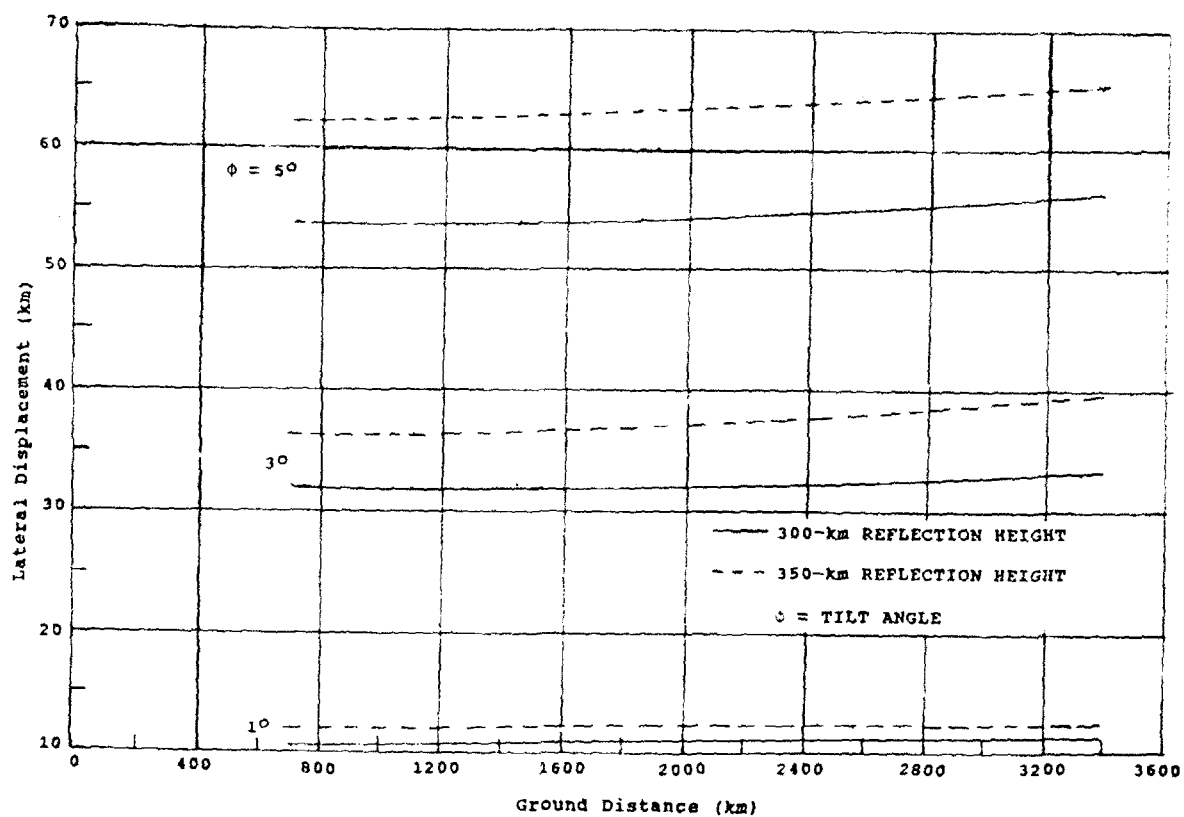


Figure 8. Lateral Displacement Due to Tilt at Horizontal Rotation of  $90^\circ$

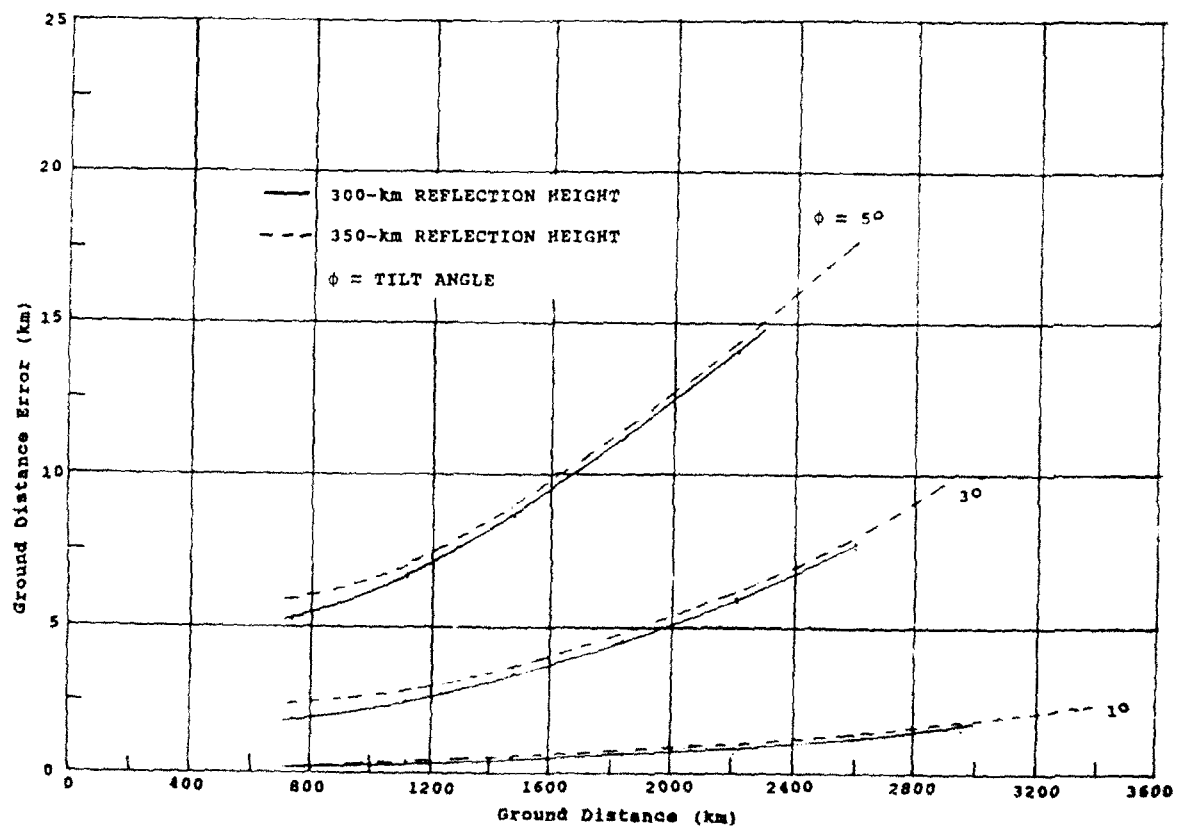


Figure 9. Ground Distance Error Due to Tilt at Horizontal Rotation of  $0^\circ$  and  $180^\circ$

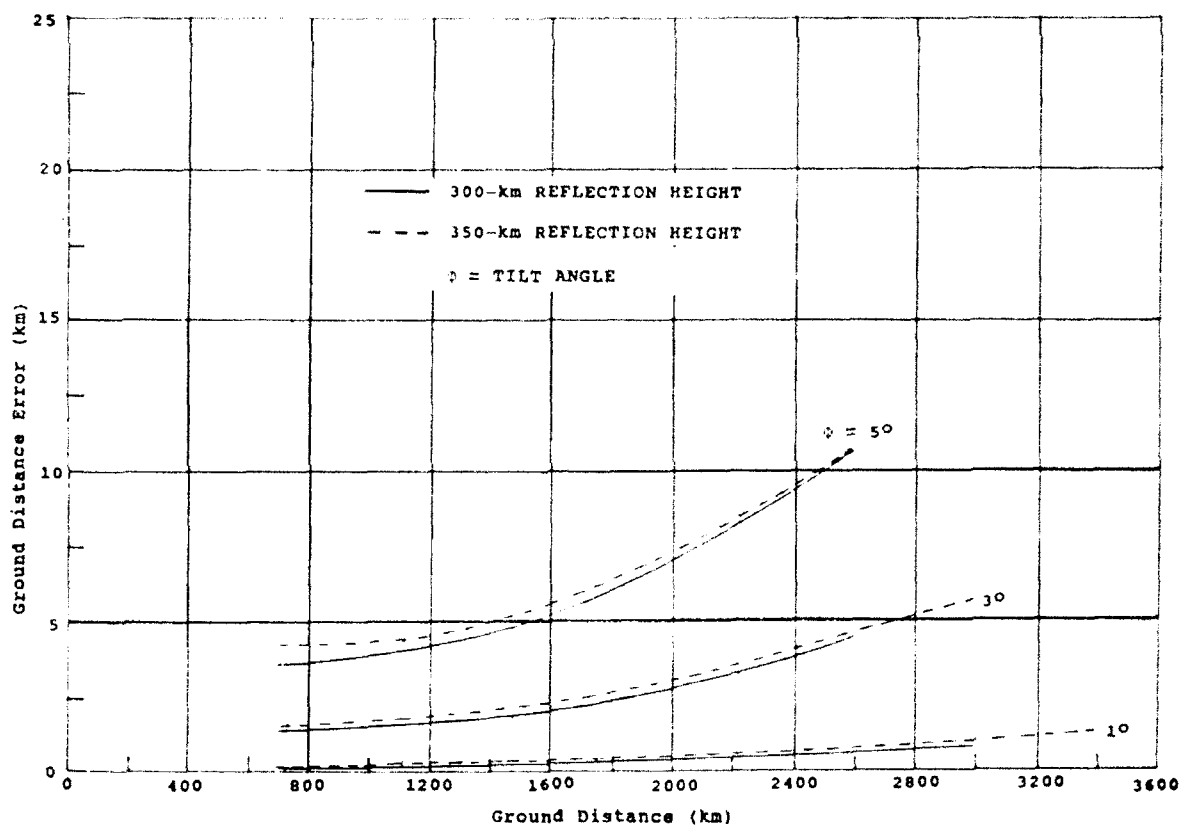


Figure 10. Ground Distance Error Due to Tilt at Horizontal Rotation of  $45^\circ$  and  $135^\circ$

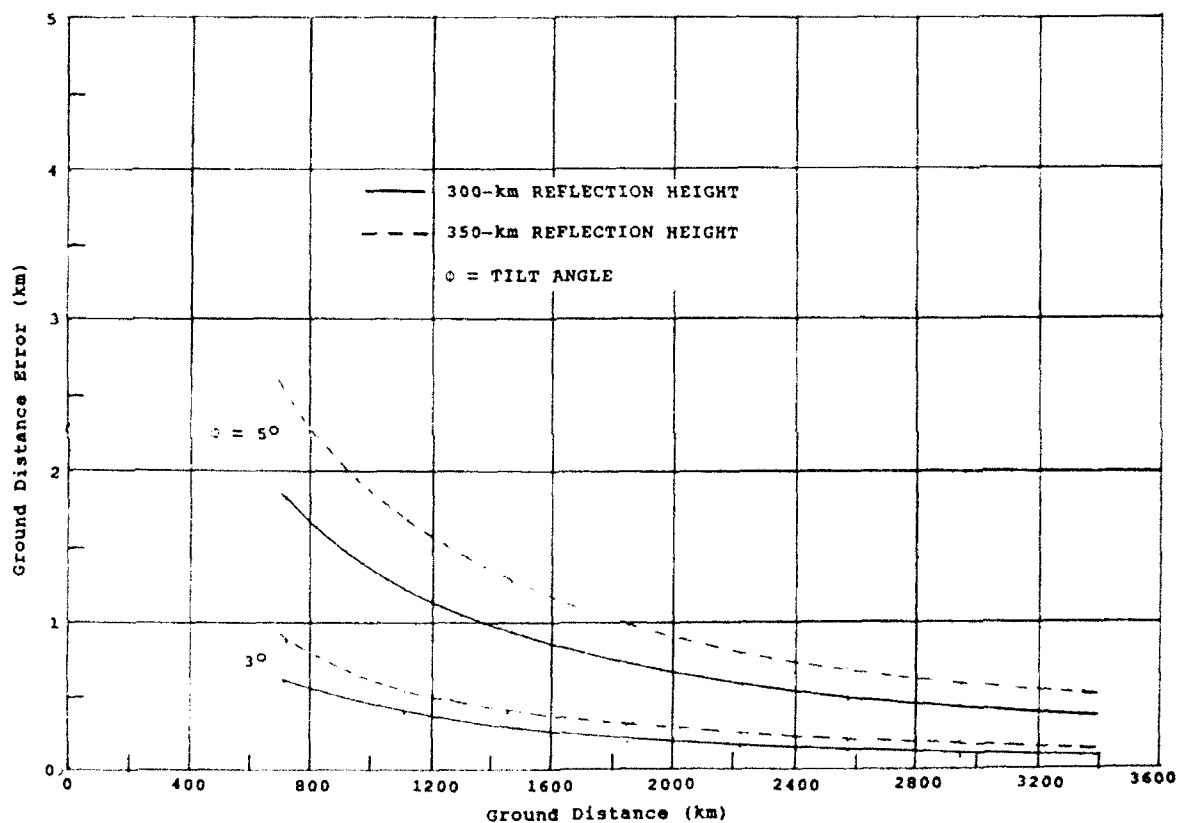


Figure 11. Ground Distance Error Due to Tilt at Horizontal Rotation of  $90^\circ$

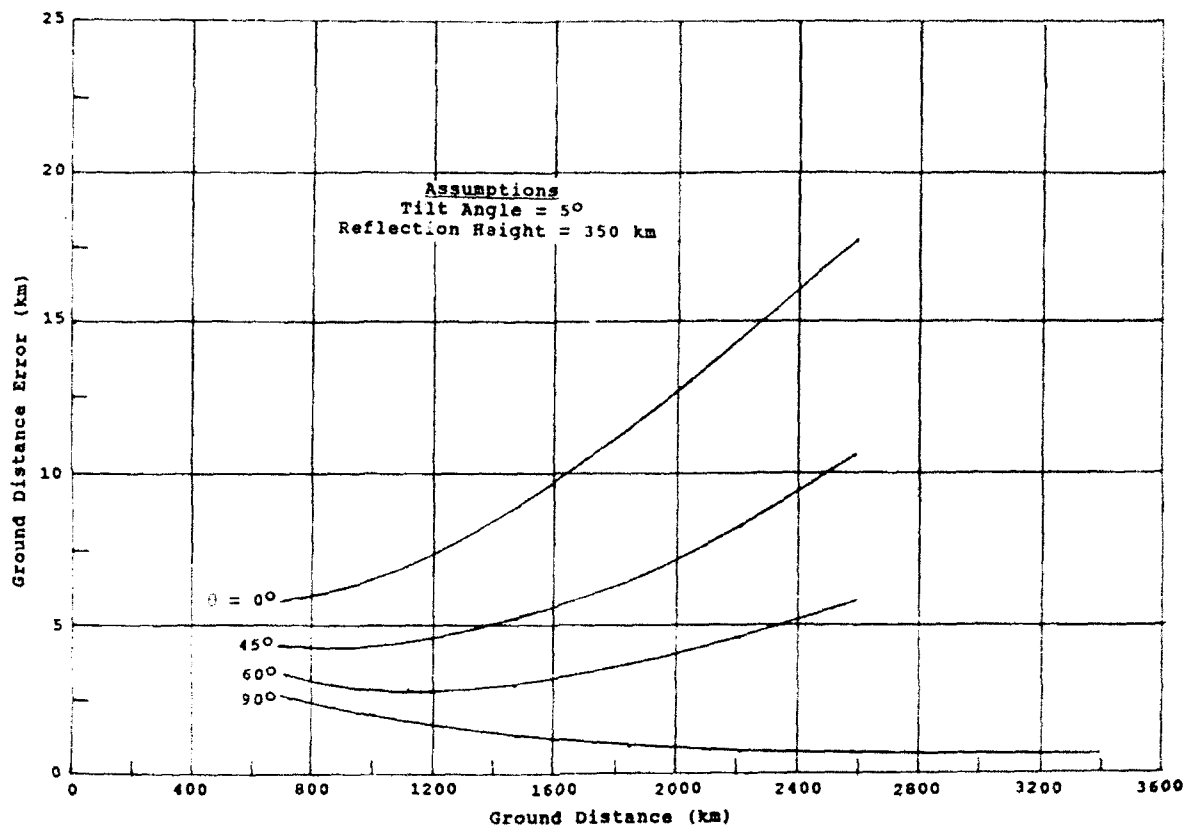


Figure 12. Ground Distance Error as a Function of Horizontal Rotation Angle,  $\theta$

## DISCUSSION

### R. ROSE

COMMENT: I noted that the ground distance error data that were presented reflected target ranges of greater than 1000 km. It should be noted that for ranges of less than 1000 km, the ground distance errors would increase dramatically.

### R. OTT

Do you have any position error results for the case where the frequency is greater than the critical frequency so just the ground wave is present?

### AUTHOR'S REPLY

This particular condition was not considered in this analysis.

### C. GOUTELARD

L'introduction de composantes transverses du gradient horizontal, souvent négligé, me paraît importante. Cependant, les déviations que vous trouvez sont faibles. Nous trouvons par des tracés de rayons des écarts qui peuvent atteindre plus de 150 km, notamment pour le rayon haut. Cela est rendu dans la communication que nous présentons ("Nouvelles méthodes de modélisation.. "[5]). Comment cette différence peut-elle être expliquée?

*The introduction of transverse components across a horizontal gradient - an aspect often neglected - seems to me to be important. Nevertheless, the deviations you found are small. Through ray tracing we find deviations of more than 150 km, in particular for the high ray. This is explained in the paper we are presenting ("New Methods of Modeling"). How can this difference be explained?*

### AUTHOR'S REPLY

The maximum tilt angle and reflection height assumed in this analysis are  $5^\circ$  and 350 km, respectively. Larger tilt angles and higher reflection heights would yield ground distance errors and lateral deviations larger than those presented in this paper. An analysis of plasma frequencies along its  $75^\circ$  west meridian reveals that tilt angles larger than  $10^\circ$  are possible. Electron density gradients at sunrise and sunset can also produce tilt angles in the order of  $10^\circ$ .

# NOUVELLE METHODE DE MODELISATION MESO-ECHELLE DE L'IONOSPHERE A PARTIR D'UN SITE UNIQUE. APPLICATION A LA RADIOLOCALISATION H.F.

## NEW METHOD OF MODELLING THE MESOSCALE STRUCTURE OF THE IONOSPHERE WITH A SINGLE SENSOR. IMPACT ON H.F. RADIO-LOCATION

J.L. MOKRZYCKI, L. BARTHES, J. CARATORI, C. GOUTELARD

LETTI  
Université Paris-Sud  
Bâtiment 214  
91405 Orsay Cedex  
France

### I. - GENERALITES -

La modélisation de l'ionosphère est une nécessité dans de nombreuses applications : étude physique du milieu, télécommunications, radars transhorizon et radiolocalisation.

La radiolocalisation à partir d'un site unique est particulièrement attractive mais impose, dans le domaine des propagations décimétriques, de connaître avec une précision suffisante le milieu ionosphérique dans lequel les trajectoires des rayons sont contrôlées, non seulement par le profil vertical de l'ionisation, mais également par les gradients horizontaux.

La figure 1 représente un tracé de rayon dans une ionosphère définie par un modèle de Bradley-Dudeney [1] en présence de gradients horizontaux dont les composantes transversales ne sont pas nulles.

On peut voir que ces composantes entraînent des propagations hors du grand cercle avec des déviations au sol importantes lorsque la réflexion s'effectue dans la région F. La prise en compte de ces gradients est indispensable pour une localisation correcte, ce qui rend nécessaire leur estimation.

La radiolocalisation à partir d'un site unique nécessite donc de disposer en un même lieu d'un moyen de mesure. Un sondeur à rétrodiffusion met facilement en évidence les anisotropies (figure 2). Il s'agit d'un moyen simple susceptible d'être mis en oeuvre avec de très faibles puissances et donc attrayant.

La difficulté principale de cette technique réside dans la résolution d'un problème inverse qui, à partir des mesures effectuées, permet de retrouver en tout point d'une zone d'observation, le profil vertical duquel se

déduisent les gradients. Il s'agit d'un problème difficile qui n'a pas encore vraiment reçu de solution totalement satisfaisante malgré les efforts importants produits par le monde scientifique.

La stabilisation de la solution apparaît possible grâce à de nouvelles techniques qui autorisent la mesure de nouveaux paramètres. L'angle d'élévation des rayons rétrodiffusés est l'un de ces paramètres dont l'importance avait pourtant été signalée dans les années 1960.

L'utilisation de ce paramètre, conjointement aux données fournies par un sondeur à rétrodiffusion permettant une observation tout azimut, est développée dans cette présentation.

### II. - METHODES D'INVERSION DE L'IONOGRAMME DE RETRODIFFUSION -

Les modèles des milieux de propagation permettent, par des programmes de calcul, de déterminer avec précision les trajectoires des ondes électromagnétiques.

Ce problème, le problème direct, est la plupart du temps résolu avec une bonne précision.

Le problème inverse, qui consiste à chercher le modèle à partir d'un ensemble de mesures, appartient le plus souvent à la classe des problèmes "mathématiquement mal posés".

Dans le cas de l'inversion des ionogrammes de rétrodiffusion cette appartenance est due, soit au fait que la solution du problème inverse n'est pas unique, soit au fait que de faibles erreurs de mesure entraînent de fortes erreurs sur le modèle.

Les travaux sur l'inversion des ionogrammes de rétrodiffusion remontent aux années 1960. En 1968, C. Goutelard [2] a proposé une méthode prenant en compte les gradients d'ionisation horizontaux en généralisant la méthode des courbes de transmission de N. Smith.

En 1969, V.E. Hatfield [3] a proposé une méthode analytique intéressante bien que ne prenant pas en compte les gradients horizontaux. N.N. Rao [4] en 1974 et S.L. Chuang, K.C. Yeh [5] en 1976 proposent des méthodes itératives basées sur des techniques voisines de la pseudo solution mais qui ne prennent pas en compte les gradients horizontaux. R.E. Dubroff, N.N. Rao, K.C. Yeh [6] introduisent en 1978 des gradients horizontaux et constatent, comme J. Caratori et C. Goutelard [7] et J. Caratori [8] les difficultés qui apparaissent à cause, notamment, de la non unicité des solutions et de leurs instabilités. J. Caratori et C. Goutelard [7] proposent alors de stabiliser les solutions par des mesures tout azimut, ce qui a pour effet de rendre unique la solution dans la majorité des cas. L.E. Bertel, D.G. Cole et R. Fleury [9] introduisent en 1988 des informations supplémentaires provenant de la connaissance du diagramme de rayonnement des antennes et J.Y. Le Huerou [10] en 1989 établit une méthode permettant, par des sondages panoramiques, de déterminer les directions des gradients.

Le souci de stabiliser les solutions par l'introduction de l'angle d'élévation a été énoncé dès les années 1968 par C. Goutelard [2] et 1969 par V.E. Hatfield [3] mais les techniques, malgré les essais qui ont été effectués à l'époque, ne permettaient pas de réaliser des mesures suffisamment fiables.

A partir de 1985, deux projets français de systèmes à rétrodiffusion introduisaient la mesure des angles d'élévation : le projet de radar transhorizon NOSTRADAMUS [11] et le projet de sondeur du C.N.E.T. à l'île de Losquet [12]. Des études prenaient en compte la mesure de cet angle. L.E. Bertel, D.G. Cole, R. Fleury [9], J.L. Mokrzycki [13] et mettaient en évidence l'importance de ce paramètre dans la stabilisation des solutions.

Cet article rapporte une partie des études menées au LETTI sur ce sujet.

### III. - IMPORTANCE DE LA MESURE DE L'ANGLE D'ELEVATION -

La trace frontale de l'ionogramme de rétrodiffusion a été le support principal des méthodes d'inversion. Le temps de groupe qui lui correspond est facile à mesurer, même s'il ne l'est pas toujours avec une

précision suffisante, et le volume des mesures correspondantes est réduit.

A ce temps de groupe on peut ajouter, on l'a cité, une exploration panoramique qui permet de stabiliser les solutions.

L'adjonction de la mesure de l'angle d'élévation pour améliorer la méthode, apparaît parmi les plus simples : le paramètre s'ajoute simplement aux mesures précédentes et les équations de propagation l'introduisent naturellement.

La difficulté est plutôt expérimentale et réside dans la précision avec laquelle la mesure est effectuée. Une bonne précision nécessite de grands réseaux d'antennes et l'influence du bruit est fondamentale sur la précision des mesures. On peut noter que ces paramètres sont liés et que l'augmentation du bruit exige, pour une précision constante, un accroissement des dimensions du réseau d'antennes.

Le LETTI (Laboratoire d'ETude des Transmissions Ionosphériques de l'Université Paris-Sud) a choisi d'utiliser des traitements de signaux appropriés pour échapper à des dimensions importantes des réseaux d'antennes.

Dès lors que la précision sur l'angle d'élévation  $\beta$  est suffisante - de l'ordre de  $\pm 1$  degré - l'inversion peut être effectuée dans de bonnes conditions.

L'apport de la mesure de l'angle d'élévation peut être examiné simplement à partir des courbes de transmission généralisées, décrites dans [2], dans lesquelles on suppose une inclinaison globale de l'ionosphère E (figures 3a et 3b). Ces courbes s'utilisent comme dans la méthode proposée par N. Smith, mais elles sont tracées pour des temps de propagation constants, et non pour des distances constantes, pour des angles d'incidence constants et pour des inclinaisons de l'ionosphère E différentes de 0 degré. La seule mesure du temps de focalisation définit une courbe unique tangente à l'ionogramme zénithal. L'adjonction de l'angle d'élévation (figure 3c) fixe le point de tangence, en donne directement accès à l'ionogramme, donc au profil.

La méthode proposée s'inspire de cette remarque, en la généralisant, et développe les calculs sous forme analytique.

### IV. - METHODE DU LETTI -

Le paragraphe précédent a souligné la nécessité d'introduire une variable supplémentaire pour stabiliser le processus d'inversion. Cette variable est l'angle d'élévation  $\beta$ . Les données du problème

consistent donc en un ensemble de points  $(\beta, Pg)$ , représentant l'échantillonnée de la courbe  $Pg(\beta)$  à fréquence et azimut de sondage constants.

Comme pour les méthodes précédemment développées au LETTI [7] [8] [10] [14] [15], on admet qu'un sondage zénithal donne le profil vertical, ou "modèle d'ordre 0", et on adopte les hypothèses simplificatrices suivantes :

- L'ionosphère est représentée par un modèle vertical de Bradley- Dudeney [1].
- Le facteur de forme de la couche F2 :  $Ryh = ymF2/hmF2$  est supposé constant dans toute la zone explorée, et prend la valeur mesurée en zénithal au centre de celle-ci.
- Seules la fréquence critique  $Fc = foF2$  et la hauteur du maximum d'ionisation  $Hm = hmF2$  varient avec la distance  $D = R\theta$  par rapport au centre de la zone.
- Les gradients qui résultent de ces variations sont supposés linéaires. Ils définissent un "modèle d'ordre 1", pour lequel  $Fc$  et  $Hm$  suivent les lois suivantes :

$$\begin{aligned} Fc(D) &= Fco (1 + Gf \cdot D) \\ Hm(D) &= Hmo (1 + Gh \cdot D) \end{aligned} \quad (1)$$

où  $Gf$  et  $Gh$  sont les gradients relatifs exprimés en  $km^{-1}$ .

- La région E est supposée n'introduire que des variations faibles sur les rayons réfléchis dans la région F et une estimation de ces variations permet de corriger leurs effets.

Dans ces conditions, les paramètres à mesurer sont ceux de la région F2, et l'énoncé du problème inverse devient le suivant : déterminer les intensités  $Gf$  et  $Gh$  des gradients, connaissant l'échantillonnée de la courbe  $Pg(\beta)$  mesurée par rétrodiffusion.

Le premier point à vérifier concerne l'unicité de la solution. Celui-ci a été analysé par la comparaison entre elles des courbes  $Pg(\beta)$  simulées pour l'ensemble des valeurs possibles des gradients, soit :

$$Gf, Gh \in [-3 \cdot 10^{-4} km^{-1} ; +3 \cdot 10^{-4} km^{-1}] \quad (2)$$

L'étude de la distance euclidienne entre une courbe quelconque  $Pgo(\beta)$ , associée à un couple  $(Gfo, Gho)$  donné, et toutes les autres courbes du domaine précédemment défini, a confirmé l'hypothèse d'unicité. Ce résultat peut encore s'exprimer en disant que, pour un modèle d'ordre 0 donné, et une fréquence

de sondage donnée, il n'y a pas deux courbes  $Pg(\beta)$  identiques. Notons que ce point est d'une extrême importance, et constitue un progrès considérable par rapport aux méthodes antérieures, qui ne faisaient appel qu'à la trace frontale de l'ionogramme de rétrodiffusion.

Afin de mesurer les gradients, il est nécessaire de caractériser la courbe  $Pg(\beta)$  obtenue expérimentalement. De nombreuses tentatives ont été faites, qui ont abouti à la conclusion suivante : ce sont les coordonnées  $(\beta_f, Pg_f)$  du minimum de la courbe  $Pg(\beta)$  qui assurent la caractérisation la plus efficace de cette courbe. On pourrait penser alors qu'il suffit de mesurer directement les coordonnées de ce minimum, et non la courbe complète. Un tel raisonnement ne prend pas en compte les incertitudes de mesure qui peuvent entraîner des erreurs importantes, surtout dans le cas de la détermination d'un point unique. Cette difficulté a été résolue en considérant tous les points disponibles, et en ajustant ceux-ci par un polynôme des moindres carrés de degré faible (2 à 3), de façon à lisser le bruit de mesure. Les coordonnées du minimum le plus probable sont alors obtenues en égalant à zéro la dérivée du polynôme d'interpolation.

Le second et le troisième point à vérifier concernent la bijectivité et la continuité de la relation liant les coordonnées du minimum  $(\beta_f, Pg_f)$  aux intensités des gradients  $(Gf, Gh)$ . Ici encore, la simulation de l'ensemble des cas pratiques a montré que, non seulement la relation cherchée est bijective, mais qu'elle est aussi continue. Les figures 4 et 5 illustrent ces résultats.

La première (figure 4), qui a l'aspect d'une grille, représente indifféremment les deux réseaux de courbes à deux paramètres :

$$\begin{aligned} P'f &= \ln(Pg_f) = f(\beta_f, Gf, Gh) \\ \beta_f &= g(P'f, Gf, Gh) \end{aligned} \quad (3)$$

tracés pour un triplet  $(Hm, Ryh, x = F/Fc)$  fixé. Il est clair que les courbes obtenues sont continues et ne présentent aucun point double.

Les figures 5a et 5b montrent respectivement comment se déforme la grille précédente lorsque  $Hm$ , ou  $Ryh$ , ou  $x$  varient. On constate que les remarques précédentes restent valables pour toutes les valeurs usuelles des paramètres concernés.

L'existence et la stabilité de la solution étant assurées, il reste à modéliser la grille  $(\beta_f, P'f)$  en fonction des paramètres  $Gf, Gh, Hm$  et  $Ryh$  à fréquence réduite  $x$  constante.

Dans ce but, nous avons commencé par ajuster la "croix centrale" de chaque grille par deux polynômes des moindres carrés d'ordre 2 à 2 variables. Ceux-ci s'écrivent :

(4)

$$\begin{pmatrix} P'f \\ \beta_f \end{pmatrix} = (A) \begin{pmatrix} Gf^2 \\ Gh^2 \end{pmatrix} + (B) \begin{pmatrix} Gf \\ Gh \end{pmatrix} + \begin{pmatrix} P'fo \\ \beta_{fo} \end{pmatrix}$$

avec :

(A), (B) matrices carrées d'ordre 2.

$(P'fo, \beta_{fo})^T$  vecteur définissant le centre de la croix.

Chacun des 10 coefficients  $p_k$  de ces équations a été à son tour ajusté par un polynôme des moindres carrés d'ordre 4 à 2 variables, de la forme :

$$p_k = \sum_{i=0}^4 \sum_{j=0}^4 \sigma_{kij} Hm^i Ryh^j \quad (5)$$

A l'issue de ce processus d'ajustement, toutes les croix centrales correspondant à une fréquence réduite  $x$  donnée, sont représentées par un ensemble de 250 coefficients  $\sigma_{kij}$ .

Le système d'équations 4 peut être utilisé directement pour réaliser l'inversion, plus précisément pour fournir une valeur approchée  $(Gfi, Ghi)$  des gradients, en fonction des coordonnées  $(\beta_f, P'f)$  du minimum de la courbe  $Pg(\beta)$ . La figure 6 montre sur un exemple l'écart qui existe entre les solutions exactes  $(Gf, Gh)$ , représentées par les points d'intersection du quadrillage en trait plein, et les solutions approchées  $(Gfi, Ghi)$ , représentées par de petits cercles. Ce résultat s'explique simplement si on remarque que l'équation 4 représente en toute rigueur la croix centrale de la grille de la figure 4, et que son utilisation pour reconstituer la grille complète, conduit simplement à translater selon  $Gf$ , ou selon  $Gh$ , l'une des branches de la croix centrale. Or, la figure 4 montre que la grille réelle ne se déduit pas de la croix centrale par simple translation. Une modification de la méthode d'inversion est donc nécessaire.

En pratique, il est apparu qu'il était beaucoup plus simple de corriger après coup les valeurs  $(Gfi, Ghi)$  issues des formules 4, plutôt que compliquer celles-ci en modélisant les grilles par des polynômes d'ordre supérieur à deux. Nous avons alors écrit :

$$\begin{aligned} Gfc &= Gfi + \delta Gf \\ Ghc &= Ghi + \delta Gh \end{aligned} \quad (6)$$

où

$Gfc, Ghc$  sont les valeurs corrigées des gradients.

$\delta Gf, \delta Gh$  sont les écarts calculés.

Les fonctions écart ont ensuite été ajustées par des polynômes des moindres carrés de la forme :

$$\delta G_{h,f} = \sum_{i=0}^4 \sum_{j=0}^4 c_{ij} Gh^i Gf^j \quad (7)$$

expression dans laquelle les  $c_{ij}$  sont eux-mêmes des fonctions de  $Hm$  et de  $Ryh$  données par :

$$c_{ij} = \sum_{m=0}^2 \sum_{n=0}^2 d_{ijmn} Hm^m Ryh^n \quad (8)$$

L'ensemble des formules de correction précédentes utilise un total de 450 coefficients  $d_{ijmn}$ . La précision obtenue, ou précision intrinsèque de la méthode, est alors excellente. La figure 7 illustre cette proposition. On y trouve, comme sur la figure 6, les solutions exactes et les solutions approchées, mais aussi les solutions corrigées  $(Gfc, Ghc)$  représentées par des croix. Il apparaît que ces dernières sont indiscernables des solutions exactes. Le calcul montre que l'écart type de l'erreur résiduelle est compris entre 4 et  $5 \cdot 10^{-8} \text{ km}^{-1}$  pour  $Gf$  comme pour  $Gh$ . Comme les plus petits gradients que nous cherchons à déterminer sont de l'ordre de  $10^{-5} \text{ km}^{-1}$ , l'incertitude due aux formules d'inversion est d'environ 0,5%, ce qui est amplement suffisant.

## V. - EVALUATION DE LA PRECISION GLOBALE -

La figure 8 représente l'organigramme de la méthode d'inversion développée dans le paragraphe précédent. Un sondage zénithal fournit le modèle d'ordre 0. La fréquence critique  $F_c$  permet de fixer la fréquence du sondage par rétrodiffusion à élévation variable, soit  $F_e = 1,5F_c$  dans cet exemple. Les échantillons  $(\beta, Pg)$  de la courbe  $Pg(\beta)$  sont interpolés, et les coordonnées  $(\beta_f, P'f)$  du minimum sont calculées. Les formules d'inversion 4 et 5 permettent d'obtenir la solution approchée  $(Gfi, Ghi)$ , puis enfin la solution corrigée grâce aux relations 7 et 8.

Nous avons montré que la précision intrinsèque du processus d'inversion est excellente (0,5%). En pratique cependant, les grandeurs mesurées sont entachées d'erreurs, lesquelles sont introduites dans les formules d'inversion et entraînent donc des incertitudes sur les valeurs des résultats. Il est donc indispensable de compléter l'étude précédente par une évaluation de la précision globale, laquelle caractérise la "robustesse" de l'ensemble du processus d'inversion dans des conditions d'exploitation réelles.

Pour réaliser cette étude, nous sommes partis de la courbe  $Pg(\beta)$  théorique, e. nous avons simulé les



incertitudes de mesure en ajoutant aux coordonnées de chaque point de cette courbe une erreur aléatoire à distribution uniforme (figure 9). Nous avons posé :

$$\begin{aligned} P_{gm} &= P_{gt} + \delta P_g \\ \beta_{gm} &= \beta_t + \delta \beta \end{aligned} \quad (9)$$

où

$P_{gt}, \beta_t$  sont les valeurs théoriques  
 $P_{gm}, \beta_{gm}$  sont les valeurs "mesurées"  
 $\delta P_g, \delta \beta$  sont les incertitudes de mesure.

La figure 10 est un exemple de courbe "mesurée", simulée par ce procédé.

Le processus d'inversion est ensuite appliqué à la courbe précédente, et les gradients "mesurés" ( $G_{fm}, G_{hm}$ ) sont comparés aux gradients théoriques ( $G_{ft}, G_{ht}$ ).

L'ensemble de ces opérations a été appliqué une centaine de fois à chacun des modèles ayant servi à l'étude, de manière à obtenir une statistique de l'incertitude globale. De plus, quatre classes d'erreur de mesure ont été considérées, correspondant aux couples ( $\delta P_m, \delta \beta_m$ ) suivants : (50km, 2°), (50km, 1°), (25km, 2°) et (25km, 1°). Le tableau de la figure 11 donne les résultats obtenus dans 2 cas extrêmes.

On observe que la précision est meilleure pour  $H_m$  grand, ce qui concorde avec le fait que la "grille" est alors plus étalée, les mailles sont plus larges, donc la sensibilité aux variations de  $P_g$  et  $\beta$  est plus faible (voir figure 5a). D'autre part, les erreurs sont du même ordre de grandeur pour les deux gradients, et elles décroissent plus vite lorsque  $\delta \beta_m$  diminue, que lorsque c'est  $\delta P_m$  qui diminue. Cela signifie qu'en pratique, il faudra réaliser des réseaux d'antennes ayant une excellente résolution angulaire pour pouvoir atteindre une précision de l'ordre de  $10^{-5} \text{ km}^{-1}$ , mais aujourd'hui cela ne constitue plus une barrière technologique.

Quoi qu'il en soit, même avec une résolution de 50km sur  $P_g$  et de 2° sur  $\beta$ , la méthode proposée améliore la précision d'un facteur au moins égal à 5 par rapport aux méthodes antérieures, ce qui constitue un progrès remarquable.

Signalons enfin que la méthode d'inversion a été développée pour déterminer les paramètres de la couche F2, mais que l'adjonction des paramètres de la couche E suffit à définir le modèle de Bradley-Dudeney. Les couches inférieures E et F1 introduisent essentiellement un retard de propagation et un allongement de la trajectoire, tous deux susceptibles d'être corrigés.

## VI. - CONCLUSION -

Le problème de l'inversion des ionogrammes de rétrodiffusion est à la base de nombreuses applications : télécommunications, radars transhorizon, études géophysiques et localisation à partir d'un site unique.

L'utilisation d'un sondeur ionosphérique monostatique à rétrodiffusion offre un moyen de mesure simple, efficace et peu polluant compte tenu des faibles puissances qu'il est possible d'utiliser.

Les précisions que l'on peut atteindre sur la mesure de l'angle d'élévation ( $\approx \pm 1$  à  $\pm 2^\circ$ ) permettent de stabiliser les solutions du processus d'inversion, et de déterminer les paramètres ionosphériques avec une précision supérieure - d'un facteur 5 à 10 - aux méthodes jusqu'ici utilisées.

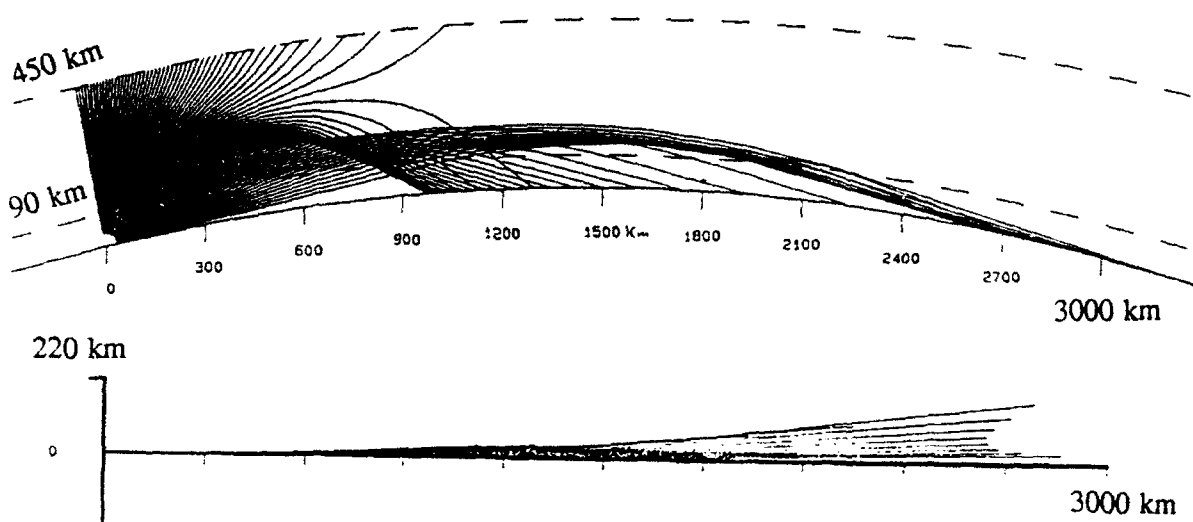
L'exploration azimutale permet de tracer une cartographie de l'ionosphère sur une zone de l'ordre de 6000 km de diamètre comme le montre la figure 12.

L'utilisation de ces cartes est précieuse dans toutes les applications, notamment dans les problèmes de localisation à partir d'un site unique.

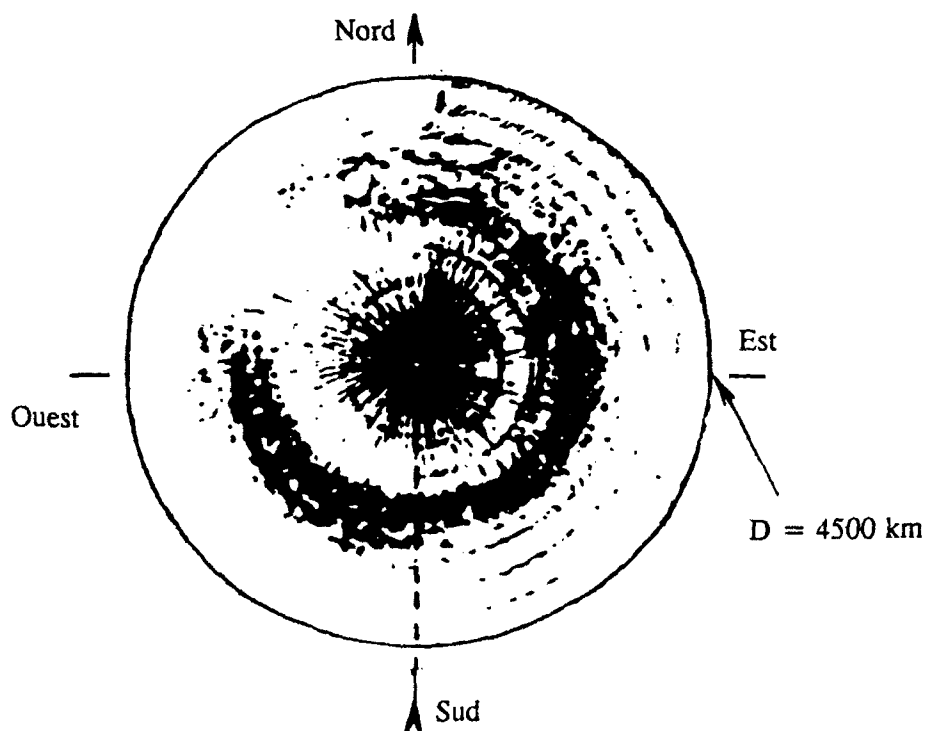
## BIBLIOGRAPHIE

- [1] P.A. BRADLEY, J.R. DUDENEY - A simple model of the vertical distribution of electron concentration in the ionosphere. J.A.T.P., Vol. 35, pp 2131-2146, 1973.
- [2] C. GOUTELARD - Analyse de la structure fine des échos de sol obtenus par rétrodiffusion des ondes décimétriques - Application à l'étude des perturbations ionosphériques. Thèse d'Etat, Paris 1968.
- [3] V.E. HATFIELD - Derivation of ionospheric parameters from backscatter data. Agard, XV meeting, Canada, sept. 1969.
- [4] N.N. RAO - Inversion of sweep frequency sky wave backscatter leading edge for quasiparabolic ionospheric layer parameters. Radio Science, Vol. 9, n° 10, pp 845-847, oct. 1974.
- [5] S.L. CHUANG, K.C. YEH - A method for inverting oblique sounding data in the ionosphere. Radio Science, Vol. 12, n° 1, pp 135-140, jan. 1977.

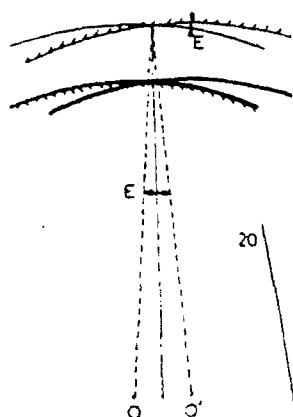
- [6] R.E. DUBROFF, N.N. RAO, K.C. YEH - Backscatter inversion in spherically asymmetric ionosphere. Radio Science, Vol. 14, n° 5, pp 837-841, sept. 1979.
- [7] J. CARATORI, C. GOUTELARD - Modélisation de l'ionosphère dans les problèmes de gestion de réseaux de transmission H.F. Congrès AGARD, Ottawa (Canada), avril 1978.
- [8] J. CARATORI - Contribution à l'étude des prévisions à très court terme des transmissions par voie ionosphérique. Détermination rapide des conditions moyennes de propagation à partir de sondages par rétrodiffusion. Thèse d'Etat, Paris, 1987.
- [9] L.E. BERTEL, D.G. COLE, R. FLEURY - The inversion of backscatter ionograms. I.P.S., IPS-TR-88-03, Sydney, 1988.
- [10] J.Y. LE HUEROU - Contribution à la modélisation de l'ionosphère. Evaluation des gradients horizontaux à partir de sondages panoramiques par rétrodiffusion. Mémoire d'ingénieur CNAM, Paris, 1989.
- [11] C. GOUTELARD - NOSTRADAMUS : Projet français de radar transhorizon. 47ème congrès AGARD EPP, Grèce, Oct. 1990.
- [12] N. RUELLE, F. GAUTHIER - Inversion of ionospheric sounding data from an HF backscatter radar. Meeting on inverse problems Conf. Proc., Springer-Verlag, 1990.
- [13] J.L. MOKRZYCKI - Modélisation de l'ionosphère à partir de sondages par rétrodiffusion à azimut et élévation variable. Mémoire d'ingénieur CNAM, Paris, Décembre 1991.
- [14] R. SCHWAB, J. CARATORI, C. GOUTELARD - Real time large scale modelling of the ionosphere with backscatter soundings. New results. Congrès IEE, Londres (G.B.), avril 1988.
- [15] J. CARATORI, C. GOUTELARD - Prévisions à très court terme par modélisation du brouillage et de l'ionosphère à l'aide de la rétrodiffusion. Congrès AGARD, San-Diego (U.S.A.), mai 1989.



**FIGURE 1 :** Tracé de rayon en présence de gradients horizontaux.  
 Modèle de Bradley-Dudeney.  
 $h_m F2 = 300 \text{ km}$ ,  $y_m F2 = 150 \text{ km}$ ,  $x = F/foF2 = 4$   
 $G_h = 1.10^{-4} \text{ km}^{-1}$ ,  $G_f = 0$ ,  $\Phi_h = 60^\circ$

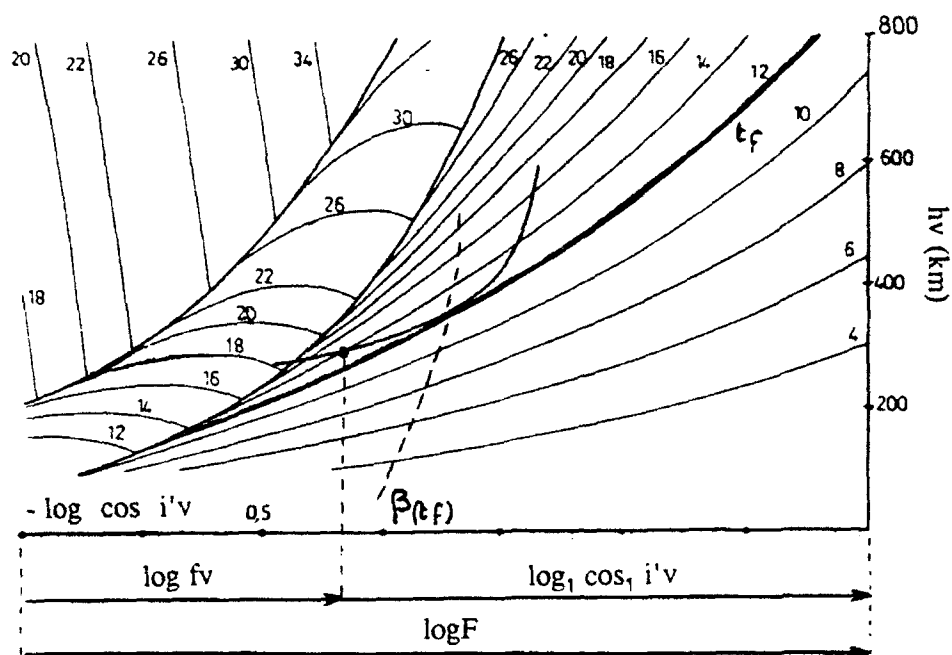


**FIGURE 2 :** Sondage par rétrodiffusion à fréquence fixe et azimuth variable.

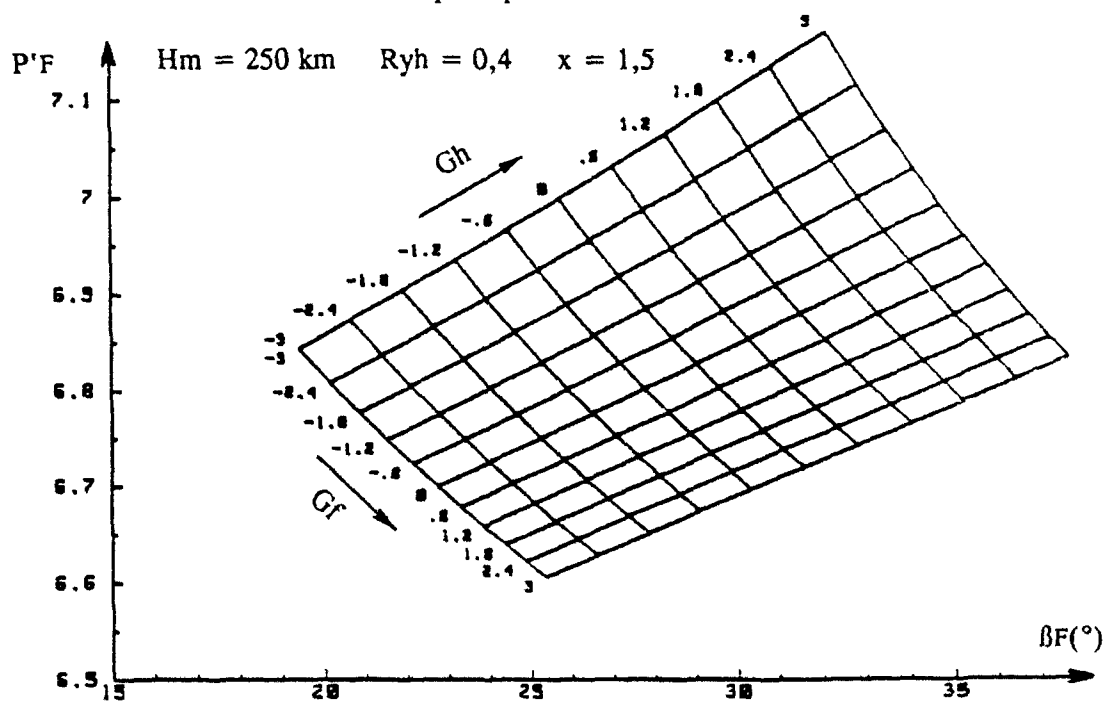


**Figure 3a**  
Définition de  
l'inclinaison E

**Figure 3b**  
Utilisation des courbes  
de transmission. Paramètre  $T_g(\text{ms})$   
Exemple pour  $E = 3^\circ$



**FIGURE 3 :** Emploi des courbes de transmission généralisées pour évaluer l'inclinaison des couches ionosphériques.



**FIGURE 4 :** Variation des coordonnées ( $BF$ ,  $P'F$ ) du point de focalisation en fonction des gradients ( $Gf$ ,  $Gh$ ), à fréquence constante.

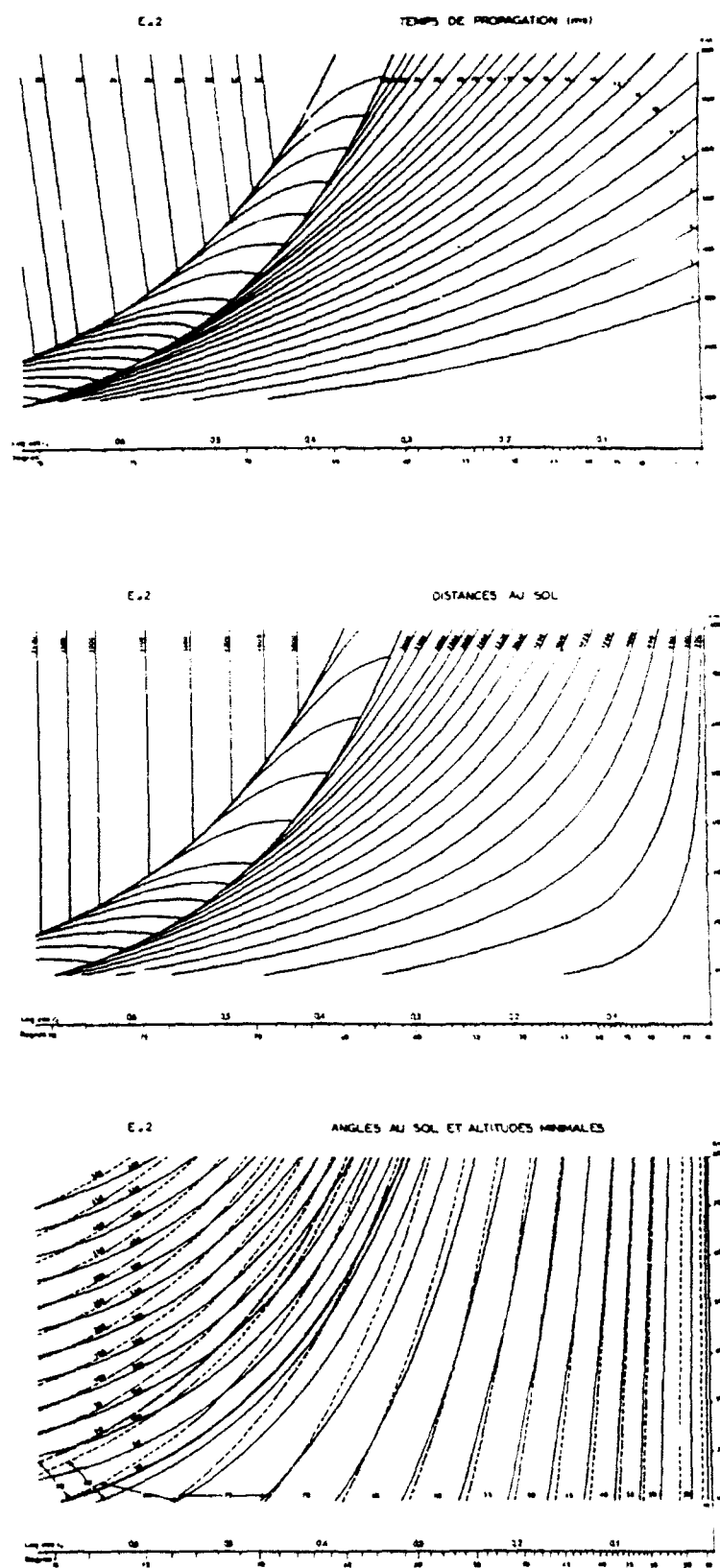
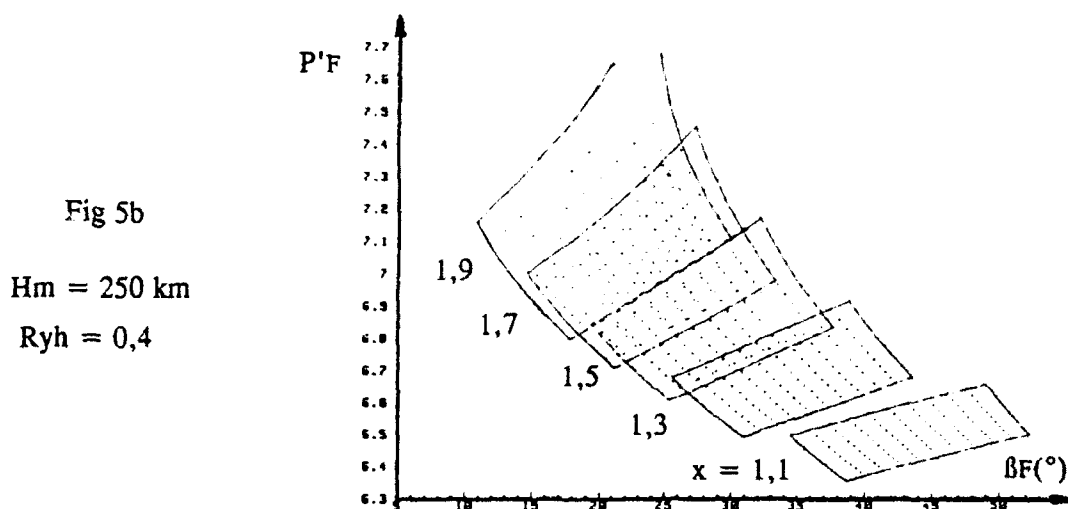
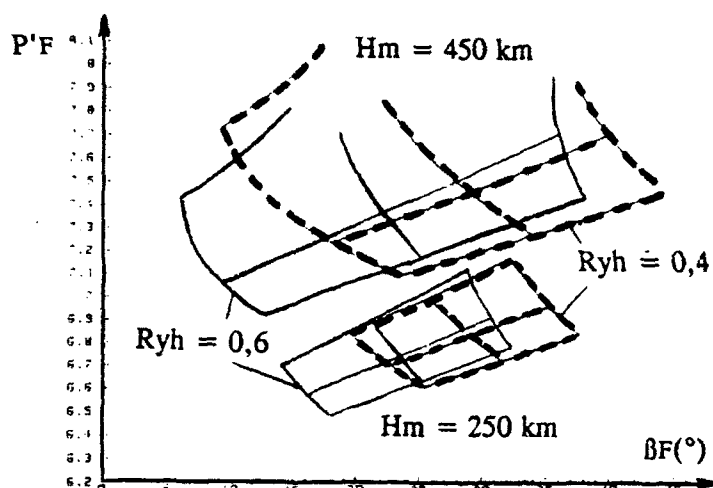
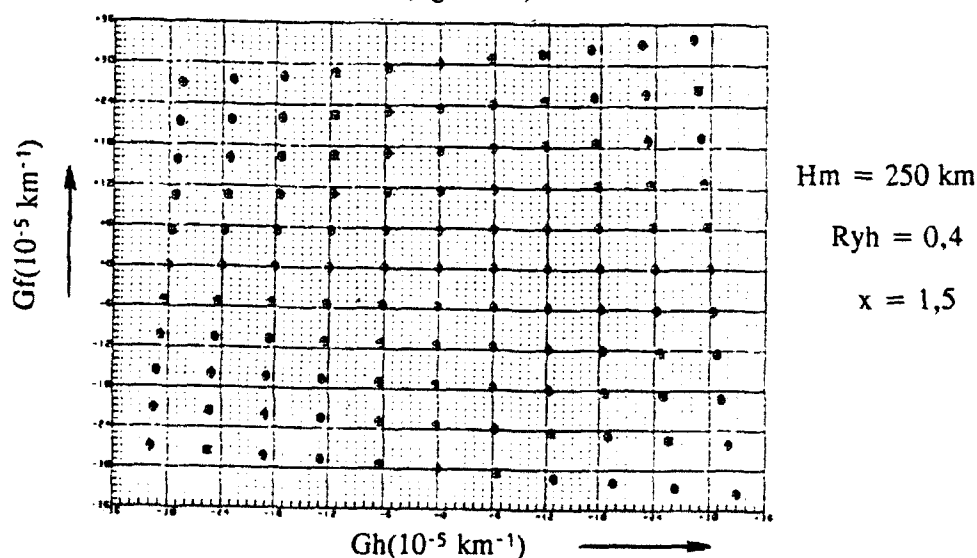


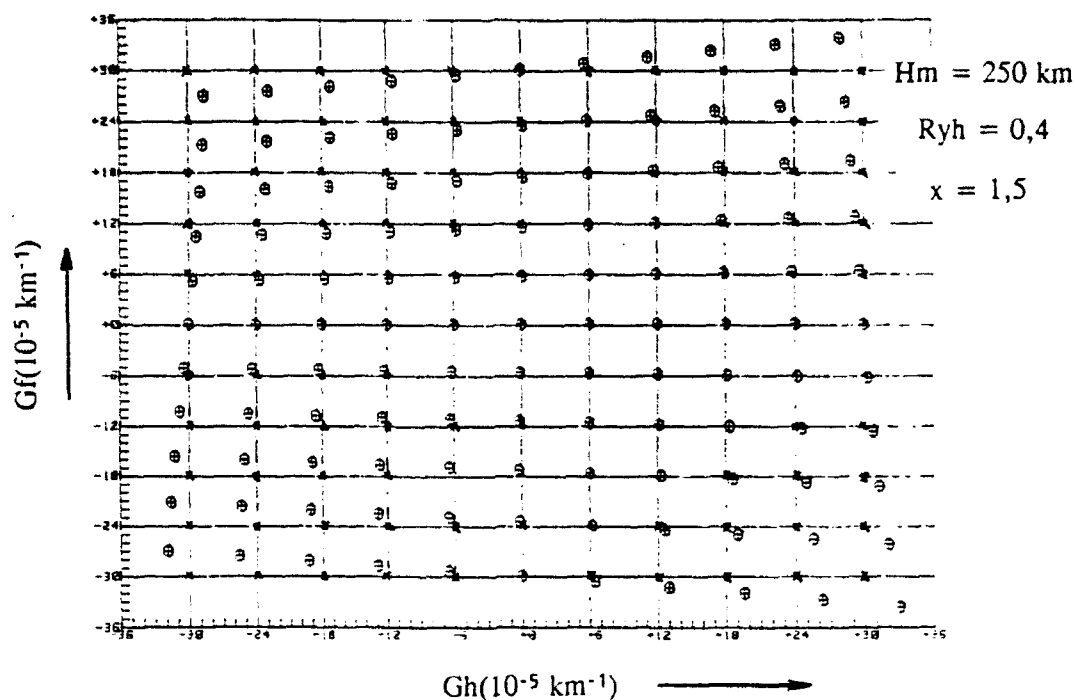
Figure 3c : Courbes de transmission généralisées pour une ionosphère en présence d'un gradient de 2 degrés.



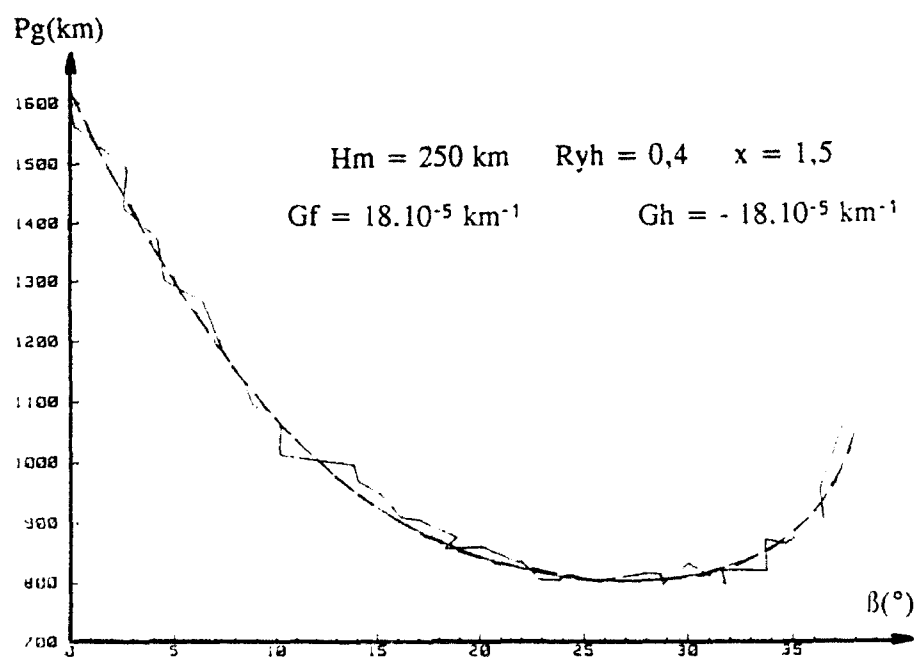
**FIGURE 5 :** Evolution de la grille  $P'F(\beta F)$  en fonction de  $H_m$  et de  $R_{yh}$  (figure 5a), et en fonction de  $x = F/F_c$  (figure 5b).



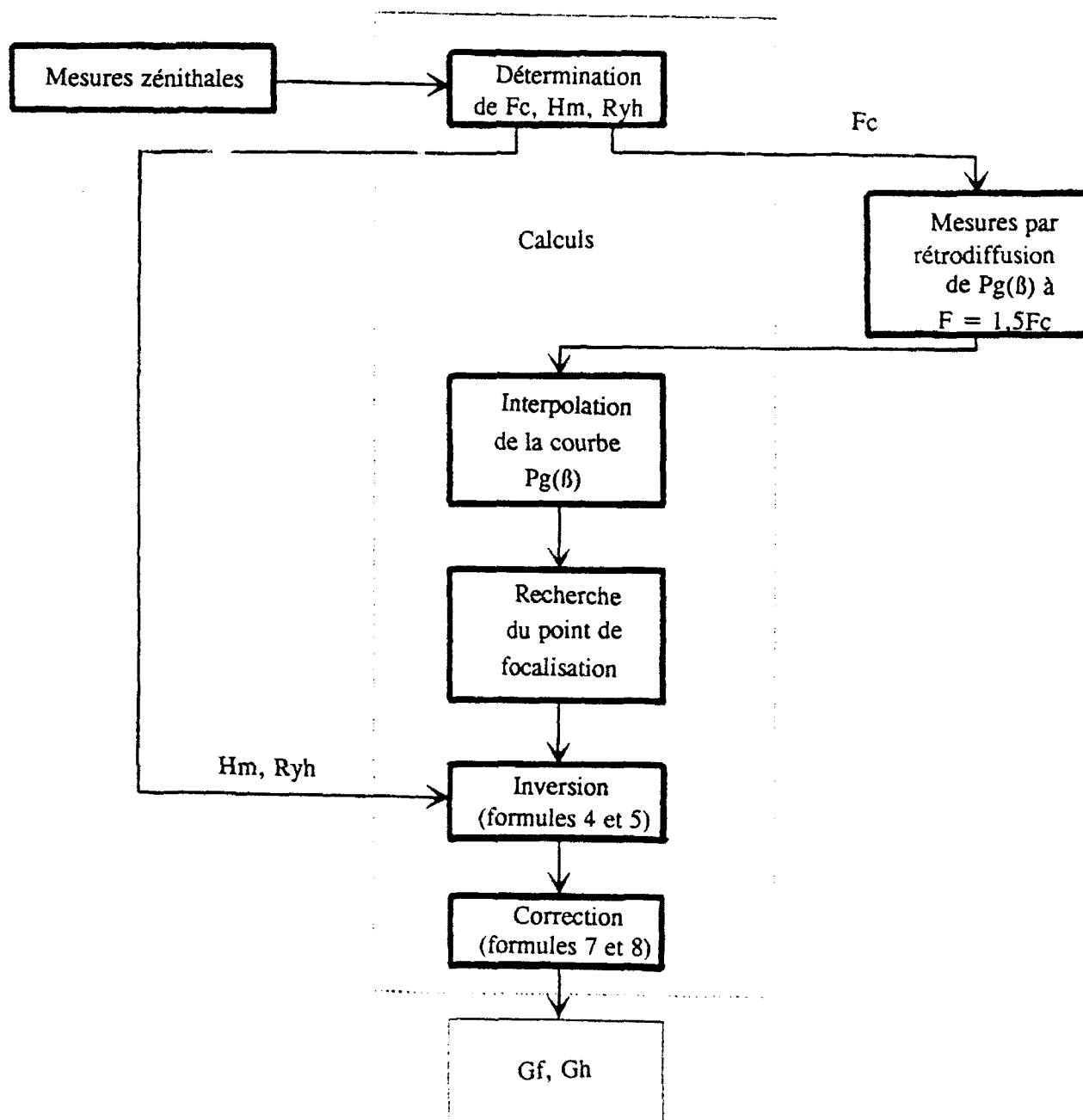
**FIGURE 6 :** Inversion sans correction : comparaison entre les solutions approchées (petits cercles) et les solutions exactes (intersections du réseau).



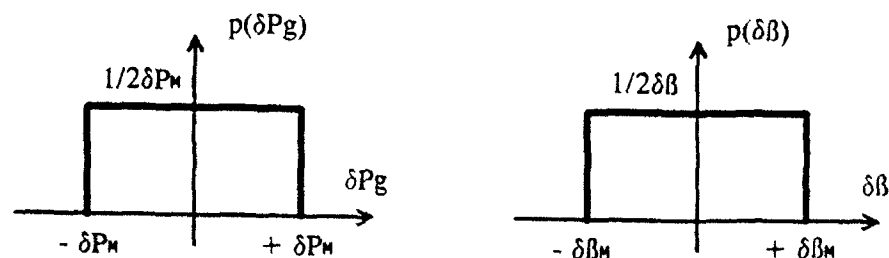
**FIGURE 7 :** Inversion avec correction : comparaison entre les solutions approchées (petits cercles), les solutions exactes (intersections du réseau), et les solutions corrigées (croix).



**FIGURE 10 :** Simulation d'une courbe  $Pg(\beta)$  "mesurée".



**FIGURE 8 :** Organigramme du processus d'inversion.

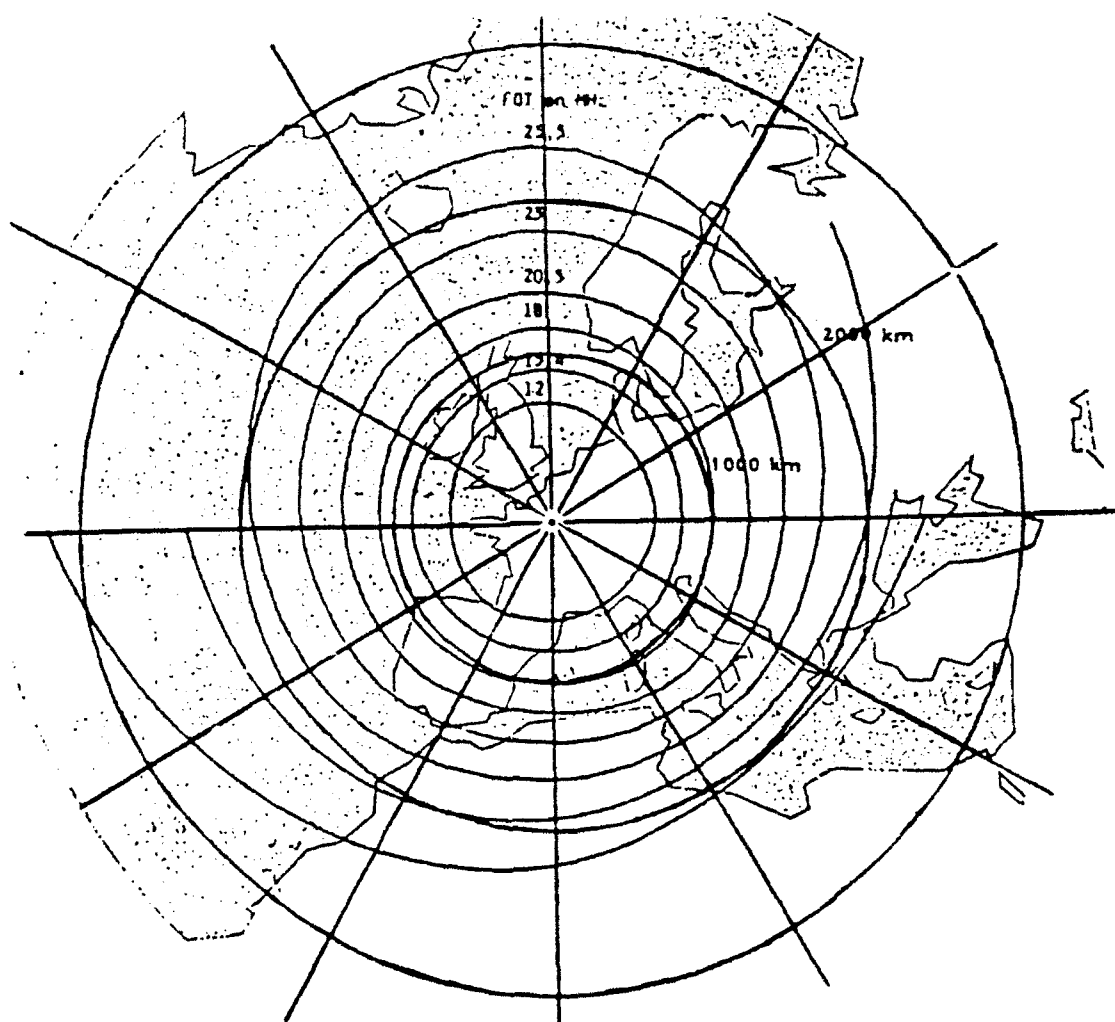


**FIGURE 9 :** Distribution des erreurs de mesure pour  $P_g$  et  $\beta$ .



Erreur absolue moyenne	$ \delta G_h  (10^{-5} \text{ km}^{-1})$				$ \delta G_f  (10^{-5} \text{ km}^{-1})$			
$\delta P_m$ (km)	50		25		50		25	
$\delta \beta_m$ (°)	2	1	2	1	2	1	2	1
Cas 1 $H_m = 450 \text{ km}$ $R_{yh} = 0,6$	1,43	0,71	1,18	0,51	1,30	0,84	1,27	0,56
Cas 2 $H_m = 250 \text{ km}$ $R_{yh} = 0,4$	2,86	2,12	2,02	1,11	2,75	2,30	1,70	1,11

**FIGURE 11 :** Précision globale de la méthode d'inversion.



**FIGURE 12 :** Cartographie de l'ionosphère montrant les courbes d'isofréquence optimale de trafic (isofot).

# THE SPATIAL AND TEMPORAL CHARACTERISTICS OF HIGH FREQUENCY AURORAL BACKSCATTER

by

D.S. Choi  
Bert Weijers  
Capt. R.J. Norris  
and N.B. Myers

Electromagnetics and Reliability Directorate  
Rome Laboratory  
Hanscom AFB, MA, 01731-5000  
United States

## ABSTRACT

Preliminary results from the measurements of high frequency (HF) backscattered signals from the auroral ionosphere using the Verona-Ava Linear Array Radar (VALAR) system are presented. VALAR is an experimental HF backscatter system capable of obtaining high resolution synoptic mapping of HF signals backscattered from field-aligned electron density irregularities in the auroral ionosphere. The receive system includes a 700 meter long linear array, providing the high azimuthal resolution required for determining the spatial distribution of HF auroral backscatter. Since the completion of the system tests and calibration at the end of 1989, experimental campaigns have been carried out on a near-monthly basis. In this paper, we provide a brief description of VALAR and present preliminary measurements of three types of phenomena: ground backscatter, slant-F, and auroral backscatter.

## 1. INTRODUCTION

HF radars operating at high latitudes are subject to periods of degradation in performance due to the presence of electron density irregularities in the auroral ionosphere. These irregularities can be envisioned as a random distribution of cylindrical blobs or patches elongated in the direction of the Earth's magnetic field. HF signals propagating in the auroral ionosphere can be strongly backscattered to the radar receiver when the wave normal is perpendicular (or almost perpendicular) to these magnetic field-aligned irregularities (Figure 1). These irregularities tend to drift with a distribution of velocities in the auroral ionosphere, causing both Doppler shifts and spreads in the backscattered signal. This unwanted backscattered signal is termed auroral clutter. The properties of auroral clutter are poorly understood, requiring extensive analysis and characterization prior to the development and evaluation of clutter mitigating techniques.

Rome Laboratory's Verona Ava Linear Array Radar (VALAR) is a unique experimental HF backscatter system dedicated to investigating the characteristics of auroral clutter. It is capable of providing the high azimuthal resolution required for determining the spatial distribution of HF auroral backscatter. VALAR also provides a unique elemental data recording capability for the development and evaluation of clutter mitigating techniques. The acquired elemental data are processed and analyzed off-line. In this paper we provide a brief description of VALAR and present

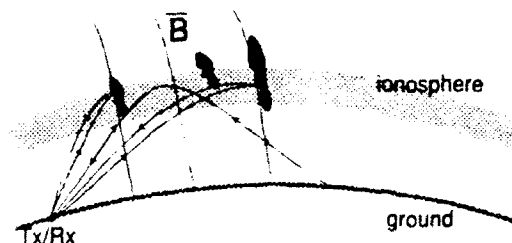


Figure 1. Ray paths of HF waves depicting direct backscatter from field-aligned irregularities in the auroral ionosphere.

some results from the preliminary analysis of data collected during 1990.

## 2. SYSTEM DESCRIPTION

VALAR was developed by Rome Laboratory to establish a dedicated experimental HF backscatter system to resolve various clutter mitigation issues. The receive system is capable of providing the narrow azimuthal beam required to accurately determine the spatial distribution of auroral clutter. The transmit system is located in Ava, NY, and the receive system is in Verona, NY. The approximate geographical location of VALAR is  $43^{\circ} 24'$  North and  $75^{\circ} 23'$  West. From this location, geographic north is about  $5^{\circ}$  east of geomagnetic north.

### 2.1 Transmit System

The Ava transmit site provides various transmitter and antenna configurations for both low and high power requirements. The facility supplies HF signals in the 2 to 30 MHz band at up to 300 kW average power. It has the capability to simultaneously transmit both a wideband Swept-Frequency Continuous Wave (SFCW) and a narrowband linear Frequency-Modulated Continuous Wave (FMCW) signal. The wideband SFCW signal is required to operate the oblique backscatter sounder located in Verona. During a typical data acquisition campaign, a north-looking rhombic antenna is used to support the low launch angle requirement of VALAR. At 10 MHz, the theoretical 3 dB elevation beamwidth of this rhombic is 14 deg centered at 10 deg elevation. The theoretical 3 dB azimuthal beamwidth is also 14 deg.

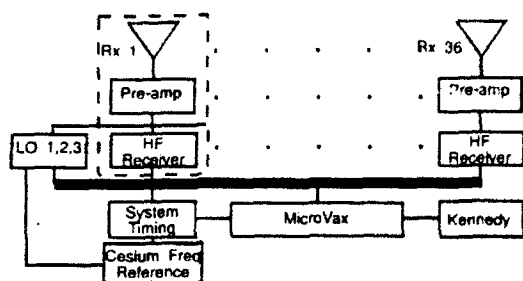
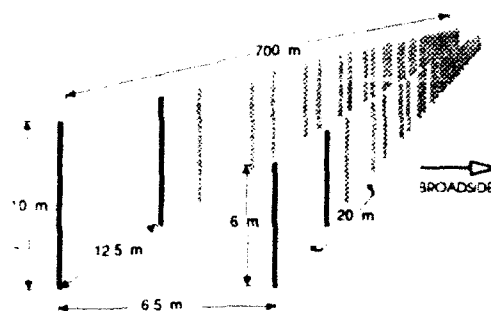


Figure 2. System block diagram. A subsystem consisting of a subarray, a preamplifier, and an HF receiver is indicated in the dotted box.

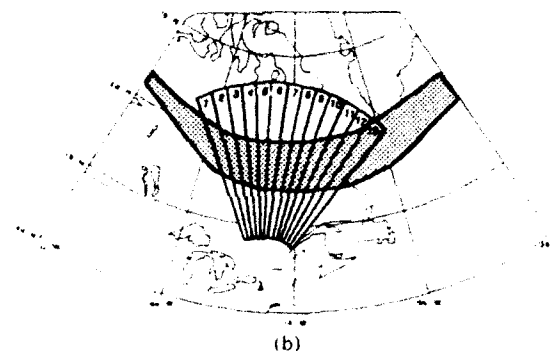
## 2.2 Receive System

A block diagram of the receive system is shown in Figure 2. The receive system consists of 36 preamplifiers, 36 identical narrow band HF receivers, a MicroVaxII, a Kennedy tape drive, and a 700 meter long linear array described below. Each receiver converts the HF analog signals to complex digital baseband data and includes a motherboard with 1 Mbyte of RAM where the digitized data are temporarily buffered prior to serial transfer to the MicroVaxII. The data acquisition software on the MicroVaxII allows the user to select various system parameters during data acquisition. In addition to the data acquisition software, diagnostic software is also available on-site. This software enables an on-line check of both system performance and data quality. The operator can simultaneously display the amplitude, phase, and signal-to-noise ratio (SNR) of all 36 channels, examine the frequency spectrum of a single channel, and plot the azimuthal distribution of the power incident on the antenna array for a given range.

The configuration of the receive array, consisting of 36 subarrays of two active and two passive monopoles, is shown in Figure 3a. The array was designed to optimize the performance over the frequency band of 6 to 12 MHz because auroral clutter is a nighttime phenomenon and the ionosphere does not support higher frequencies during this period. Previous antenna pattern measurements at 12 MHz have indicated a 2.5 deg half-power beamwidth (unweighted), 23 dB array gain, and 30 dB RMS sidelobe levels (Gould, 1990). With the boresight of the array at  $10^{\circ}$  east of geographic north, VALAR provides a coverage area extending in azimuth from  $20^{\circ}$  west to  $40^{\circ}$  east of geographic north and in slant range from 500 to 2500 km. The coverage area of VALAR is shown in Figure 3b. The 13 beams indicated in the figure are the azimuth beams used in the range-Doppler processing. These 13 beams are steered at 5 deg intervals over the 60 deg coverage of VALAR. For example, beams 1, 7, and 9 are  $20^{\circ}$  west,  $10^{\circ}$  east, and  $20^{\circ}$  east of geographic north, respectively. Also depicted in Figure 3b is the approximate position of the auroral oval at 2100 hr local time (LT) for a typical magnetic index of  $Q=3$ . Note that both the equatorward and polarward boundaries of the auroral oval are within the field of view of VALAR.



(a)



(b)

Figure 3. (a) Configuration of VALAR antenna array. The four darker elements indicated form one of 36 subarrays (b) Coverage area of VALAR and the expected position of the auroral oval at 2100 hr for magnetic index  $Q=3$ .

## 3. HF BACKSCATTER MEASUREMENTS

On September 24, 1990, HF backscatter data were collected from 1900 to 2400 hr (all times referenced hereafter will be in local time). Auroral backscatter began to appear in the ionograms around 2040 hr. A sequence of snapshots of HF backscatter data were acquired from 2020 to 2250 hr using a coherent integration time of 3.2 seconds and an operating frequency of 10.58 MHz. A waveform repetition frequency (WRF) of 25 Hz was selected to remove range ambiguities up to 6000 km, and this limited the Doppler bandwidth to 25 Hz. At 2250 hr the intensity of the auroral backscatter decreased significantly at 10.58 MHz and the operating frequency was changed to 7.87 MHz.

In this section, we present measurements made during 2020 to 2250 hr. In Section 3.1, the north-look backscatter ionograms obtained at Verona, NY will be presented to identify various types of backscattered signals observed during the evening. In Section 3.2, the characteristics of ground backscatter under normal ionospheric conditions will be described. The measurements of slant-F and auroral backscatter made during 2100 to 2250 hr will follow in Section 3.3 and 3.4, respectively. These will be presented in the form of

amplitude-range-azimuth (ARA) maps and amplitude-range-Doppler (ARD) maps. The ARA maps display the slant range and azimuth distribution of the backscattered signals prior to Doppler processing, while the ARD maps display the slant range and Doppler distribution of the backscattered signals for a given azimuth beam. For the ARA maps, the height of the ionosphere was assumed to be 300 km to overlay the HF backscatter data on a coastline map.

### 3.1 Backscatter Ionograms

The north-look backscatter ionogram obtained at 2030 hr is shown in Figure 4a. It is a typical backscatter ionogram obtained under normal propagation conditions during early evening. The returned signals at frequencies between 6.5 to 9.5 MHz are from the first and second vertical incidences. The slant range of the first vertical incidence at 6.5 MHz is about 300 km which is a typical F-layer height. As expected, the slant range of the second vertical incidence is about twice the first. Beyond 9.5 MHz, the ground backscatter emanates from the second vertical incidence and increases in slant range with increasing frequency. The ground backscatter return indicates a signal that is twice reflected from the F-layer in returning to the receiver via a reciprocal ray path. The increase in slant range with increasing frequency is expected since the oblique angle of reflection increases with increasing frequency. At 10.58 MHz, the leading edge of the ground backscatter is at 800 km.

The backscatter ionogram obtained at 2125 hr is shown in Figure 4b. The slant range of the leading edge of the ground backscatter at 10.58 MHz is approximately 1200 km. The increase in slant range from 800 km at 2030 hr to 1200 km at 2125 hr is due to the increase in F-layer height during the latter part of the night. Three additional backscattered signals are indicated in the ionogram as slant-F and auroral backscatter, respectively. At 10.58 MHz, the leading edge of the slant-F is at 500 km while the leading edges of auroral backscatters are at 800 and 1300 km in slant range. Slant-F echoes are the direct backscattered signals from field-aligned irregularities in the F-region ionosphere (Bates, 1960). From the location of VALAR, they are associated with irregularities in the subauroral F-layer (Tsunoda et al, 1981). The auroral backscatter is associated with echoes from the auroral F region and may appear at both long and short slant ranges. The long slant range is associated with high elevation angle while the short slant range implies low elevation angle (Montbrand, 1988).

### 3.2 Ground Backscatter

The range-azimuth distribution (ARA map) of the backscattered signals at 2030 and 2130 hr is shown in Figure 5a and 5b, respectively. The dominant features in Figure 5a are the bands of backscattered signals between 800 and 1200 km, uniformly distributed in azimuth. In Figure 5b, the bands of backscattered signals are visible between 1200 and 1800 km. Since the range extent of these backscattered signals are in a good agreement with the range extent of the ground

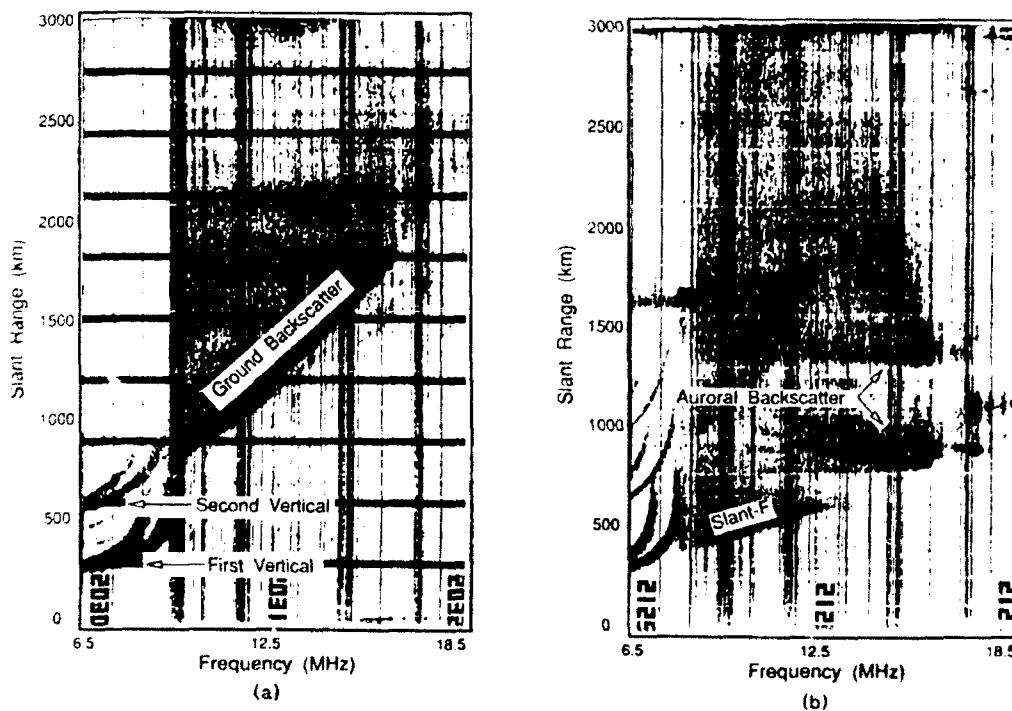


Figure 4. North-look backscatter ionograms obtained on September 24, 1990 at (a) 2030 hr, showing the ground backscatter under normal propagation condition and (b) 2125 hr, showing additional echoes, slant-F and auroral backscatter.

sep24 322

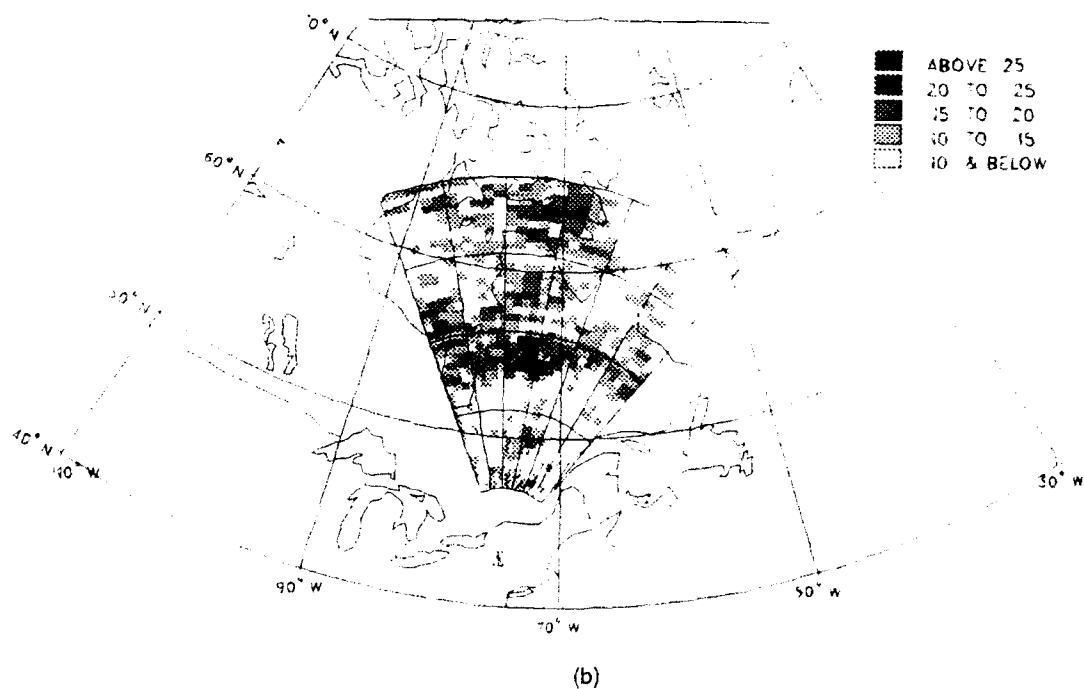
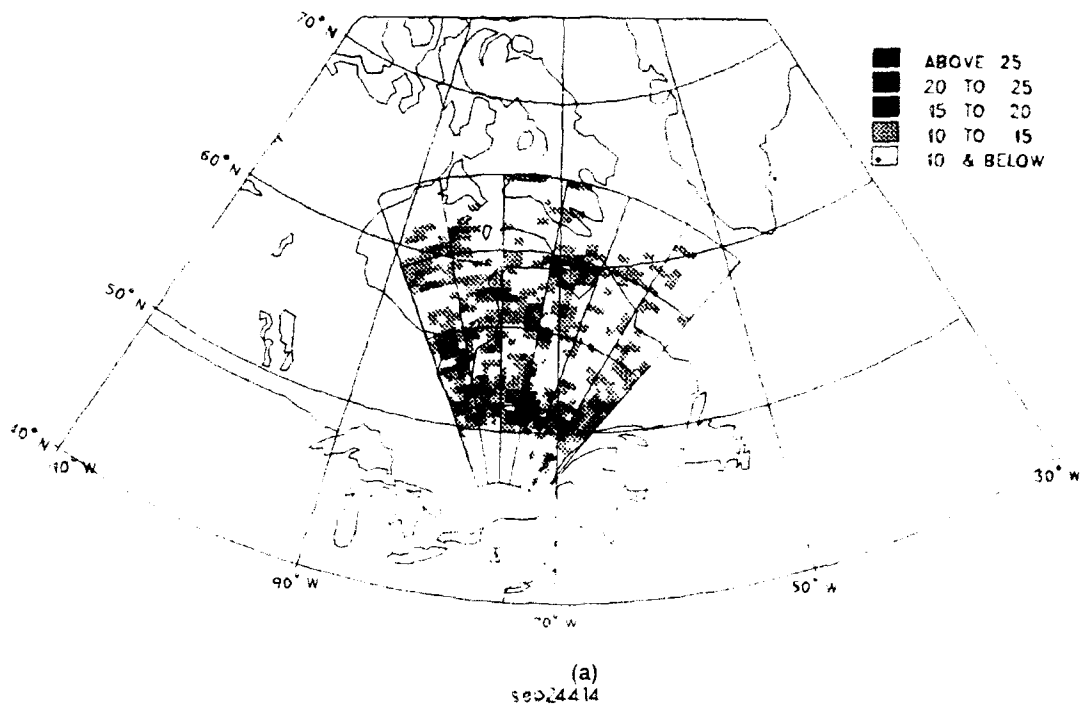


Figure 5. Amplitude-range-azimuth (ARA) maps obtained on September 24, 1990, at (a) 2030 and (b) 2130 hr. (a) shows the ground backscatter at 800 km and (b) displays the additional backscattered signals at 500, 800, and 1300 km. The contour levels are in 5dB steps.

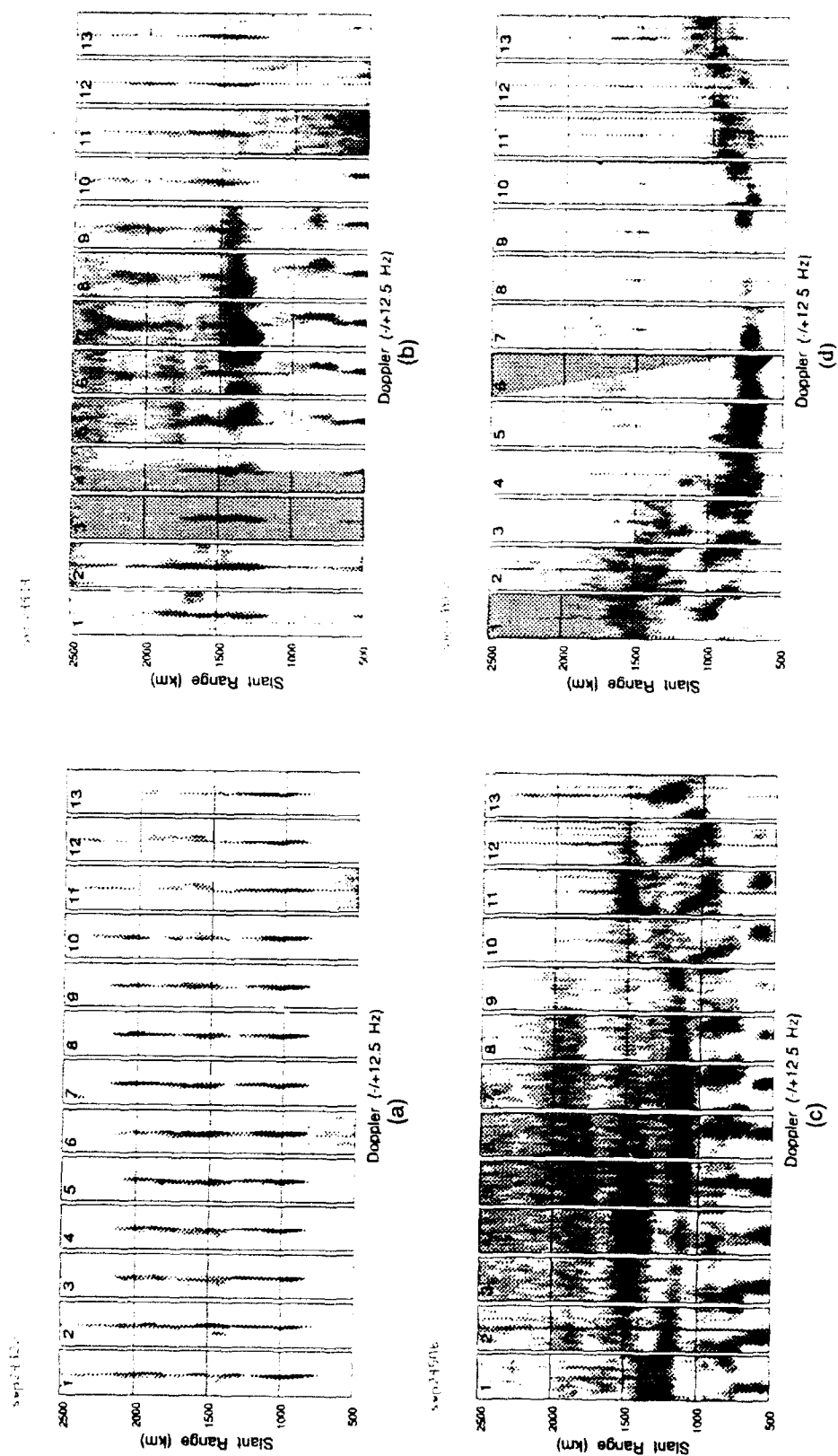


Figure 6. Amplitude-range-Doppler (ARD) maps obtained on September 24, 1990, at (a) 2030, (b) 2130, (c) 2202, and (d) 2242 hr. The azimuth beam numbers are denoted in the upper left corner of each ARD map. ARD maps in (a) show the ground backscatter exhibiting 0 Hz peak Doppler frequency. ARD maps in (b) - (d) represent the measurements of backscattered signals at early evening, peak auroral activity, and late evening period, respectively. The contour levels are in 5dB steps.

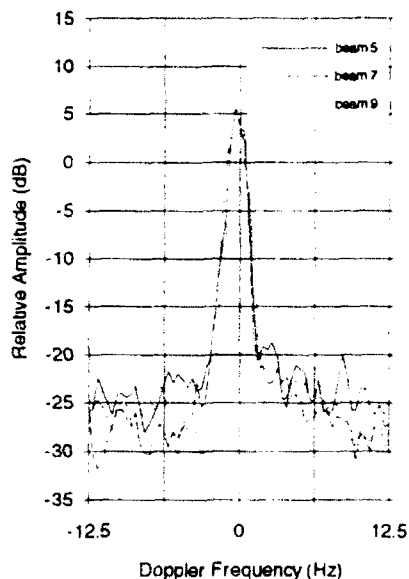


Figure 7 Doppler spectra of ground backscatter at 2030 hr for beams 5, 7, and 9.

backscatter observed in the ionograms, we can conclude that these bands of backscattered signals are ground backscatter.

We can further characterize the ground backscatter by analyzing the ARD maps obtained at these times as shown in Figure 6a and 6b, respectively. Since the ground is a stationary target, it is characterized by a peak Doppler frequency of 0 Hz. Hence ground backscatter can be identified in these ARD maps as backscattered signals with peak Doppler signature at 0 Hz.

In Figure 6a, the bands of backscattered signals between 800 and 1200 km indeed exhibit 0 Hz peak Doppler frequency, which confirms our earlier hypothesis that these backscattered signals are ground backscatter. These maps also exhibit some backscattered signals from the ground as far as 2000 km in slant range, suggesting the presence of multi-hop modes. Figure 7 shows the Doppler spectra for beams 3, 5, and 7. These Doppler spectra are obtained by averaging 5 range bins about the range bin with peak intensity. The ground clutter is approximately 30 dB above the noise floor and the spread is confined to about  $\pm 1$  Hz about 0 Hz.

The ARD maps obtained at 2130 hr (Figure 6b) show that the bands of backscattered signals between 1200 to 1800 km consist of both ground and auroral backscatter. The ground backscatter is observed in all beams, with highest intensity in beams 1 thru 4. It is interesting to note that the ground backscatter exhibits decreased intensity and range extent in beams 5 thru 10, where the slant-F and auroral backscatter are present. This is probably so because some of the transmitted rays are directly backscattered by the field-aligned irregularities and never propagate out to the ground. Future analysis using ray tracing will confirm our hypothesis.

### 3.3 Slant-F

Slant-F echoes were described in Section 3.1 as directly backscattered signals from field-aligned irregularities in the subauroral F-region ionosphere. It was shown that the leading edge of slant-F is at 500 km. The ARA map obtained at this time (Figure 5b) indeed shows some backscattered signals between 500 to 600 km slant range. The ARD maps obtained at 2130, 2202, and 2242 hr are presented in Figure 6b - 6d. These ARD maps represent the measurements during early evening, peak auroral activity, and late

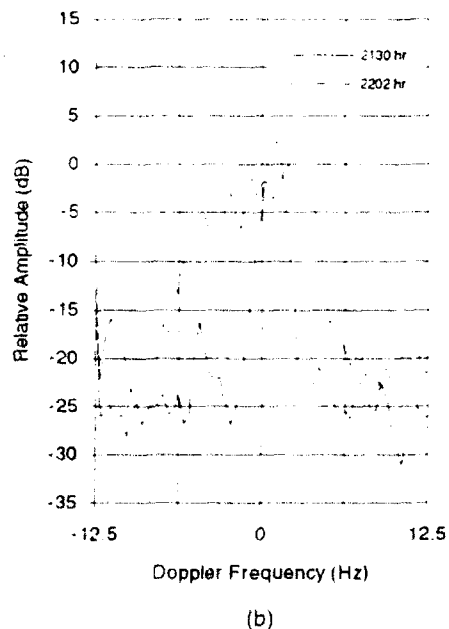
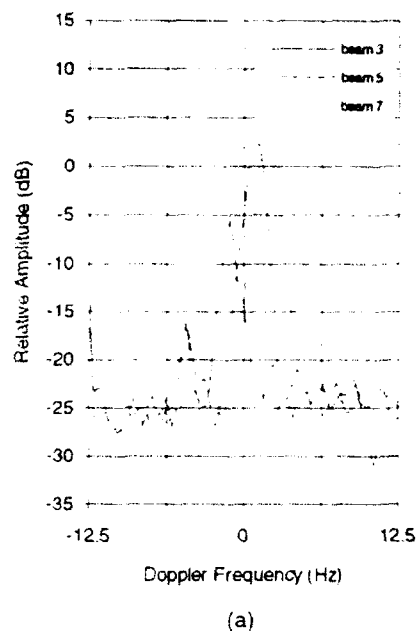


Figure 8. Doppler spectra of slant-F echoes (a) at 2130 hr for beams 3, 5, and 7 and (b) of beam 5 at 2130 and 2202 hr.

evening period, respectively. Slant-F echoes are seen in the ARD maps in Figure 6b and 6c, but they are absent in the measurement made at 2242 hr. This suggests that slant-F is a time dependent phenomenon, with peak activity between 2100 and 2200 hr on this night.

The early evening slant-F echoes exhibit Doppler characteristics resembling that of the ground backscatter. The peak Doppler spectra of slant-F echoes at 2130 hr are shown in Figure 8a for beams 3, 5, and 7. It is evident from this figure that the slant-F echoes exhibit very little spread about the peak Doppler frequency, similar to the Doppler characteristic of the ground backscatter. Unlike ground backscatter, however, the slant-F echoes exhibit shifts in the peak Doppler frequencies, with the amount of shifting depending on the azimuth beams.

After 2200 hr, the slant-F echoes measured during the subsequent peak auroral activity period also exhibited shifts in the peak Doppler frequencies. This is evident in the ARD maps obtained at 2202 hr shown in Figure 6c. In addition to the shift in the peak Doppler frequencies, the slant-F echoes at 2202 hr exhibit an increase in Doppler spread. This is evident in figure 8b, which displays the Doppler spectra of the slant-F echoes at 2130 and 2202 hr for the north-look beam.

### 3.4 Auroral Backscatter

Auroral backscatter were described in Section 3.1 as directly backscattered signals from the field-aligned irregularities in the auroral F-region. The ionogram obtained at 2125 hr (Figure 4b) clearly exhibits two distinct leading edge of auroral backscatter at 800 and 1300 km. The ARD maps at this time (Figure 6b) also show auroral backscatter starting at these ranges. The measurements of the auroral backscatter at 2202 hr (Figure 6c) exhibit an increase number of echoes of distinct range and azimuthal extents. These backscattered signals appear in bands, extending approximately 150 km in slant range and 10 to 20 degrees in azimuth. The ARD maps obtained at 2242 hr show a significant decrease in the number of auroral backscatter echoes, which is an evidence of decreased auroral activity.

We can categorize the auroral backscatter measured during this period as one of two types depending on the slant range of these echoes. Since the short slant range is associated with low elevation angle and the long slant range is associated with the high elevation angle, these will be labeled low ray and high ray, respectively. The low ray backscatter is due to the direct backscatter resulting from classical ray propagation as shown in Figure 9. If we assume a spherical Earth and an F-layer height of 300 km, the approximate slant range for a low ray path is 1200 km for a ray with 10 deg take-off angle. For a longer one-way slant range, the take-off angle must be less than 10 deg. Since the receive elements are vertical monopoles with very poor low angle coverage, we can establish that any backscattered signals beyond 1200 km cannot be due to direct backscatter resulting from a low ray path. The auroral backscatter which appears at slant range greater than 1200 km must be due to high rays. An example of such a ray path is shown in figure 9.

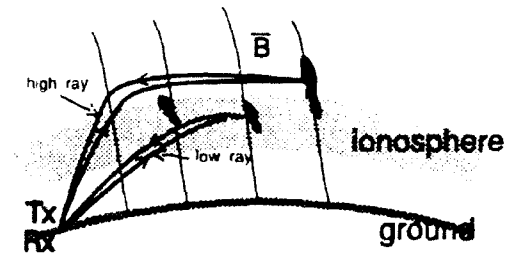


Figure 9. Paths of low ray and high ray backscattered signals from the field aligned irregularities in the auroral ionosphere.

The low ray backscatter is characterized by a shifted Doppler peak with increased intensity and spread during the peak auroral activity period. The Doppler spectra of the low ray backscatter at 800 km slant range at 2130 hrs is shown in Figure 10 for beams 5, 7, and 9. The shifts in peak Doppler frequency and the extents of Doppler spread evident in the figure are greater in comparison to those for slant-F. The spread is limited, however, suggesting a small longitudinal velocity variance.

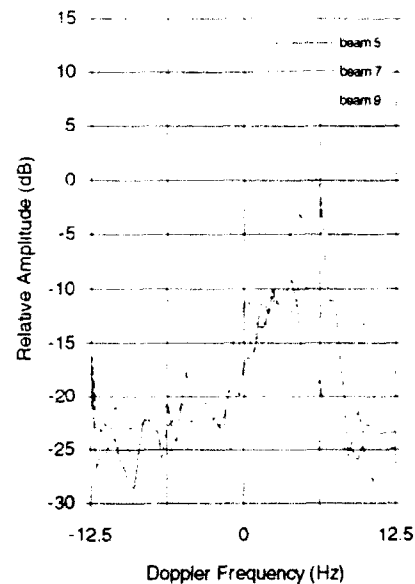


Figure 10. Doppler spectra of low ray backscattered signals at 2130 hr for beams 5, 7, and 9.

On the other hand, the high ray backscatter starting at 1250 km in Figure 6b and 1050 km in Figure 6c are characterized by high intensity and almost noise-like Doppler distribution. The peak Doppler frequency and Doppler distribution can not be inferred from these ARD maps. Figure 11a shows the Doppler spectra of the high ray backscatter at 1300 km (Figure 6b) for beams 5 and 7. The Doppler distribution of beam 5 exhibits wide frequency spread about a positive peak frequency, while the Doppler distribution of beam 7 exhibits



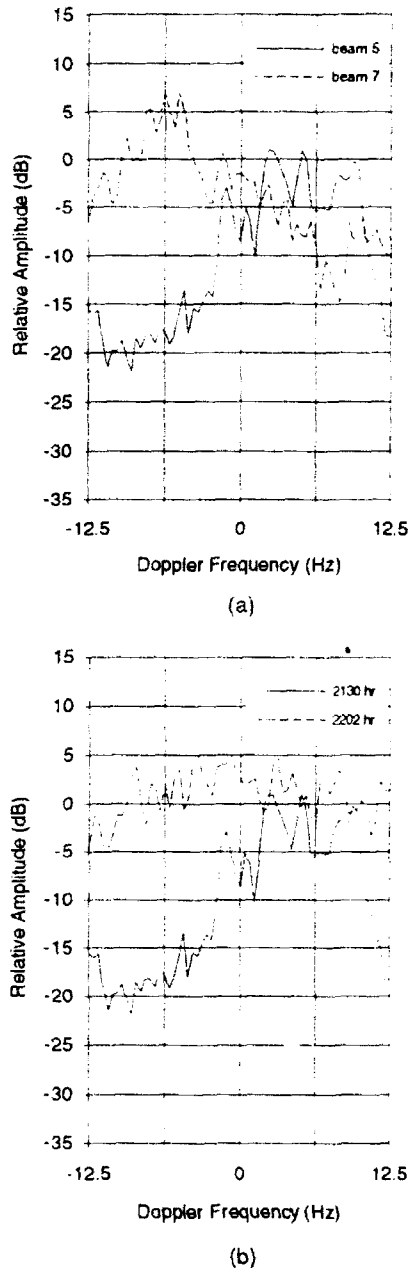


Figure 11. Doppler spectra of high ray backscattered signals (a) 2130 hr for beams 5 and 7 and (b) at 2130 and 2202 hr for beam 5.

similar spread about a negative peak frequency. Subsequently during the peak auroral activity period, the Doppler spread of the high ray signals increased even further. Figure 11b displays the Doppler spectra of the high ray signals at 2130 and 2202 hr for the north-look beam. The Doppler spectrum obtained at 2202 hr is not limited to the 25 Hz Doppler bandwidth and it resembles that of noise. The noise like characteristic exhibited by the Doppler spectrum at 2202 hr is due to the high longitudinal velocity variance of the irregularities during the peak auroral activity period.

Future data acquisition will be carried-out with a higher WRF, thereby increasing the Doppler bandwidth. It will allow us to characterize the spread extent of the Doppler and, consequently, the velocity distribution of the high ray backscatter.

#### 4. DISCUSSION

In the previous section we described the characteristics of ground backscatter and identified two types of backscattered signals from the high latitude ionosphere: slant-F and auroral backscatter. The two types have been distinguished by their origin in the ionosphere. The slant-F echoes were described as the backscattered signals from the sub-auroral F-region, while the auroral backscatter have been associated with the auroral F-region. The auroral backscatter echoes were classified as either low ray or high ray, depending on the slant range of the echoes. In this section we present a brief discussion of the data presented in section 3.3 and 3.4.

In Section 3.3 the shifts in the peak Doppler frequencies of the slant-F echoes were shown to depend on azimuth. Figure 12 displays the peak Doppler frequencies as a function of the beam steering angle for the measurements at 2130 and 2202 hr. For the azimuths west of  $-5^\circ$  (beam 4), the peak Doppler frequencies are negative, implying that the irregularities are moving away from the radar. The peak Doppler frequencies for the azimuths east of  $-5^\circ$  are positive, implying that the irregularities are moving towards the radar. This suggests that the field-aligned irregularities in the subauroral F region are moving east to west. Since the peak Doppler frequency is almost zero at beam 4, the longitudinal velocity is near zero there. Therefore, the motion of the irregularities is probably transverse to the radar look direction at this azimuth beam.

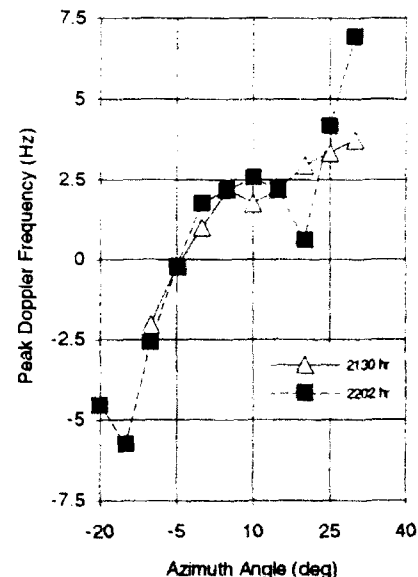


Figure 12. Peak Doppler frequency of slant-F echoes with respect to azimuth at 2130 and 2203 hr.

It is interesting to note that the shift in peak Doppler frequency is not linear with respect to azimuth. This is probably because the longitudinal velocity measured by VALAR at different azimuth beams varies as the cosine of the aspect angle, where the aspect angle is defined as the angle between the radar steering vector and the velocity vector of the irregularities.

The mean Doppler spectrum of the high ray backscattered signals obtained at 2130 hr was presented in Figure 11. It was shown that the Doppler distribution of beam 5 exhibited frequency spread about a positive peak Doppler frequency, while beam 7 exhibited a similar spread about a negative peak Doppler frequency. The Doppler characteristic of beam 5 leads to a conclusion that the irregularity structure is moving towards the radar with a large velocity variance. However, the negative Doppler distribution of beam 7, which is  $10^{\circ}$  to the east with respect to beam 5, suggest that the same irregularity is moving away from the radar, which leads to a contradiction.

This contradiction can be resolved as follows. A WRF of 25 Hz limits the Doppler bandwidth to -12.5 to 12.5 Hz, and consequently limits the maximum longitudinal velocity observable. The maximum unambiguous longitudinal velocity towards the radar for 25 Hz WRF and 10.58 MHz operating frequency is 177 m/sec. If the Doppler velocity is greater than 177 m/sec it will 'wrap-around' and appear in the negative Doppler band. Considering the expected position of the auroral oval and the direction of the electrojet current relative to VALAR at 2130 hr, the Doppler shift should be positive. Furthermore, the longitudinal velocity of the irregularity for beam 7 will be greater than beam 5. Therefore, we can conclude that the negative Doppler observed in beam 7 is due to Doppler velocities greater than 177 m/sec toward the radar.

## 5. SUMMARY

HF auroral backscatter data collected using VALAR have been presented in this paper. The capability of VALAR to observe auroral clutter makes it a unique instrument for understanding and mitigating the problems associated with HF radars operating at high latitudes. The range-azimuth distribution and Doppler characteristics of HF auroral backscatter during a two hour period were presented. Slant-F echoes were identified as originating in the sub-auroral ionosphere, and exhibited small shifts in peak Doppler frequency and very small Doppler spread. Auroral backscatter echoes were categorized as one of two types: low ray and high ray. The Doppler shifts and spreads of the low ray echoes were seen to be confined within the 25 Hz Doppler bandwidth. The Doppler shifts and spreads of the high ray echoes were shown to be much greater and not unambiguously resolved at a WRF of 25 Hz.

## 6. ACKNOWLEDGMENTS

The authors of this paper wish to thank HF system specialists John Novine and Larry Tamburrino for their invaluable support during data acquisition campaigns and to Peter Franchi and Eli Tichovolsky for their technical comments and criticisms.

## REFERENCES

- Bates, H.F. (1960), Direct HF backscatter from the F region, *J. Geophys. Res.*, 65, 1993-2002
- Gould, A. J. (1990), A thinned high frequency linear antenna array to study ionospheric structure, RADC-TR-90-186, Rome Air Development Center, Rome, NY
- Montbriand, L. E. (1988), Auroral backscatter observed at HF from Ottawa, *Radio Sci.* 23, 850-864
- Tsunoda, R. L., Basler, R. P., Showen, R. L., Walker, N. P., Frank, V. R., and Lomasney, J. M. (1981), Investigation of high-frequency radar auroral clutter and round-the-world propagation, Tech. Rep. 54, 66 pp., SRI Int., Menlo Park, Calif.

## DISCUSSION

### R. ROSE

1. What were the magnetic characteristics?
2. From the time characteristics, your measurements suggest the data show the characteristics of auroral E generated by the eastward auroral electrojet in the auroral pre-midnight sector. Have you made measurements in the post-midnight sector and does the auroral scatter have different characteristics from those you showed in your presentation?

### AUTHOR'S REPLY

1. I recall that the Kp index for September 24, 1990 was equal to 3.
2. We've had one or two data campaigns which extended until 2 a.m. local time. I recall a December 5, 1990 data set which we collected until 2 a.m. The backscattered signals showed very different characteristics and we believe that what we measured during that night was due to E layer (we have not performed mode identification, however, and we need to do that). We also would like to measure the post-midnight period sometime in the near future.

### R. JENKINS

1. Have you seen any evidence of Doppler-shifted ground clutter in your data, and if so, does it follow any consistent pattern?
2. Are the motions observed in your 'low-angle' slant-F consistent with the expected high-latitude ionospheric drift motions for the region observed?

### AUTHOR'S REPLY

1. We mostly collect data during the night-time period when ground backscatter is no longer observable so I can't say that we observed such a phenomenon.
2. Yes, the measurements tend to show that the irregularities are moving from east-to-west with respect to the geomagnetic north.

### C. GOUTELARD

*Je ferai d'abord une remarque concernant la forme de déplacement "en ovale" à laquelle vous faites allusion, il me semble que cela a été signalé par Hanuise et Villain lors d'expériences effectuées par SHERPA.*

*J'ai de plus une question qui concerne votre système. Il me semble que la rampe à 25Hz que vous utilisez introduit une ambiguïté Doppler de  $\pm 12.5\text{Hz}$  tout à fait insuffisante pour l'étude de ces perturbations pour lesquelles les Dopplers dépassent largement ces valeurs. Avez-vous prévu d'autres types de fonctionnement?*

*I would like to make an initial remark before proceeding with my question. Concerning the "oval displacement" which you refer to, I seem to remember that this was pointed out by Hanuise and Villain as part of experiments carried out by the SHERPA.*

*I also have a question concerning your system. It would seem to me that the 25Hz- ramp which you use introduces a Doppler ambiguity of  $\pm 12.5\text{Hz}$ , which is completely insufficient for study of this interference for which Dopplers easily exceed these values. Have you considered other types of operation?*

### AUTHOR'S REPLY

Yes, in fact one of the Doppler spectra shown during the presentation exhibits Doppler ambiguities. We've made changes to WRF for later data collection. Currently, we run the radar at WRF=50Hz.

# THE LONGITUDINAL OCCURRENCE OF EQUATORIAL F LAYER IRREGULARITIES

by

Jules Aarons  
Center for Space Physics  
Boston University  
725 Commonwealth Avenue  
Boston, MA. 02215  
United States

## ABSTRACT

Determining the morphology of F layer irregularities as a function of longitude in the equatorial region is vital for understanding the physics of the development of these irregularities. We aim to lay the observational basis which then can be used to test theoretical models. Theoretical models have been developed, notably in the papers by R.T. Tsunoda (Rev. Geophys. 26, 719, 1988) and by T. Maruyama and N. Matuura (J. Geophys. Res. 89, 10903, 1984). The question is whether the models are consistent with the morphology as we see it. According to our criteria, the data used should be confined to observations taken near the magnetic equator during quiet magnetic periods and at times within a few hours after sunset. Anomaly region data should be omitted for studying the generation mechanism.

The questions to be answered by proposed mechanisms are (1) why do the equinox months have high levels of occurrence over all longitudes? (2) why are there relatively high levels of occurrence in the Central Pacific Sector in the July-August period and in the 0-75° West Sector in the November-December period? (3) why are there very low levels of occurrence in November and December in the Central Pacific Sector and in July and August in the 0-75° West Sector?

Satellite in-situ data, scintillation and spread F observations will be reviewed. The limitation of each data set will be outlined particularly as relevant to the bias produced by the existence of thin versus extended layers of irregularities. A cartoon as to the occurrence pattern, as we see it, as a function of longitude will be shown.

## A. INTRODUCTION

Determining the morphology is vital for coming to grips with the physics of the development of the F-layer irregularities and the decay of these irregularities. With a knowledge of the occurrence and characteristics of F layer irregularities in the equatorial region the irregularities, mechanisms can be developed to explain the morphology as a function of longitude, for example. The aim of this paper is lay the observational basis for the development of a theory which would explain the morphology. If one could develop a definitive morphology, then theoretical models could be tested with these data. Theoretical models have been developed, notably in the papers by Tsunoda (1988) and by Maruyama and Matuura (1984). The question is whether the models explain the morphology.

While certain elements of the irregularity morphology are important for forecasting and prediction purposes for those systems that are affected by the irregularities, the data sets for the morphology to be considered in this paper are confined to those relevant to the occurrence of irregularities as a function of longitude. We shall show what elements of the total morphology of F-layer irregularities should be (and have not been) omitted. We shall evaluate the data base and give reasons for selecting data sets which are uniquely relevant to explaining the longitudinal occurrence pattern in the development of irregularities.

## B. GENERAL CHARACTERISTICS OF F-LAYER EQUATORIAL IRREGULARITIES

F-layer irregularities in the equatorial region originate at night in areas close to or on the magnetic equator. This has been determined by measuring the later time of the appearance of irregularities as they develop at latitudes away from the equator, as the disturbance rises in altitude over the magnetic equator. The later appearance of irregularities, at sites distant from the magnetic equator, indicates that the disturbance at these sites is due to an effect along the lines of force of the earth's magnetic field. The disturbance rises in altitude and affects lines of force at higher altitudes which in turn means affecting the F layer farther from the magnetic equator. To distinguish between the two distinct magnetic equator regions, some earlier papers in this field referred to the electrojet region (5° N-5° S) and the non-electrojet region i.e. from 5° to 20° North and South dip latitude. The irregularity regions which are extended in altitude include configurations that are known as plumes, patches, bubbles and blobs. In addition to the regions extended in altitude, irregularities may be concentrated in a thin layer from 200-400 km and these layers are observed only in the vicinity of the magnetic equator.

Within the concentration, the scale size of the irregularities ranges from those that are probed by radar (tens of centimeters to meters) to those that have been detected by the scintillation of satellite radio transmissions at low angles of elevation yielding a size from Fresnel considerations of the order of hundreds of meters.

It is well established that different longitudinal regions have differing annual or "seasonal" patterns of the occurrence and intensity; this will form the main area of interest of this paper. Years of high solar flux bring about both higher occurrence of equatorial irregularities and higher intensities. However we shall show that the longitudinal pattern remains approximately the same even when solar flux changes dramatically.

The effect of magnetic disturbances, using different indices as well as differing methods of analysis, has been established. Depending upon the electric field changes and the local time at the particular observatory, irregularities are either enhanced or inhibited. Since this topic is outside of the main thrust of this review we shall confine our data base, where possible, to periods of low magnetic activity.

There are various methods of making the observations. These include in-situ satellite measurements of various parameters at altitudes from 200 km to over a thousand kilometers, rocket observations of electron density, scintillation of satellite radio beacons, radar probing, and spread F detection with both bottomside and topside ionosondes. Each method measures somewhat different characteristics of the irregularities. Each biases the data in different ways. We shall try to point out the limitations of each of the methods used to detect and observe the irregularities.

We shall not consider the morphology of irregularities at the equatorial anomaly latitudes for the following reasons. For gigahertz scintillation at anomaly latitudes, we need an irregularity region at the equator extended in altitude plus a high electron density in the post-sunset period. These two requirements are not met in years of low solar flux. This may be due to the fact that electron density in the post-sunset time period may not be very high in the anomaly region in years of low solar flux. In addition to the problem of low electron density, thin layers of irregularities over the magnetic equator and layers not fully extended in altitude would not be observed in the anomaly region. The dual requirements for extended regions and the existence of thin layers mean that anomaly data cannot be used to outline the longitudinal aspects of irregularity development.

### C. RADAR

The technique that develops the most information about equatorial irregularities is radar probing of the ionosphere. The two radars which have been used are the 50 MHz radar in Jicamarca, Peru which is on the magnetic equator and the 300 MHz radar in Kwajalein Island in the Pacific which is close to the magnetic equator. Each of the radars probes for coherent scatter returns. Thus before the development of irregularities, the height of the layer and the drift motions are measured. When strong E layer irregularities develop, the returns become coherent.

Figure 1 shows radar data from Jicamarca and Figures 2 and 3 show data taken by the ALFAIR Radar in Kwajalein. For our purposes let us note the following. First in Figure 1 there is a period of time before 1929 Local Time when only a thin layer of moderate intensity has developed. In addition between the high altitude patches there are periods of intense E layer irregularities with no irregularities noted higher than 500 km altitude. It can also be seen that there is a limitation on the sensitivity of the radar in that returns are obtained, for the most part, only to 700 km. From satellite measurements, it is clear that irregularities extend to higher altitudes.

There are many forms of the irregularity structure. One aspect of importance for this study is that there are, for some evenings, thin layers of irregularities, for example from 200 to 400 km, with no evidence of irregularities at higher altitudes. For other evenings there are plumes or irregularities extending in height from 200 to over 700 km. To illustrate the importance of the height of the irregularity structure in developing a morphology of irregularities, we have utilized data

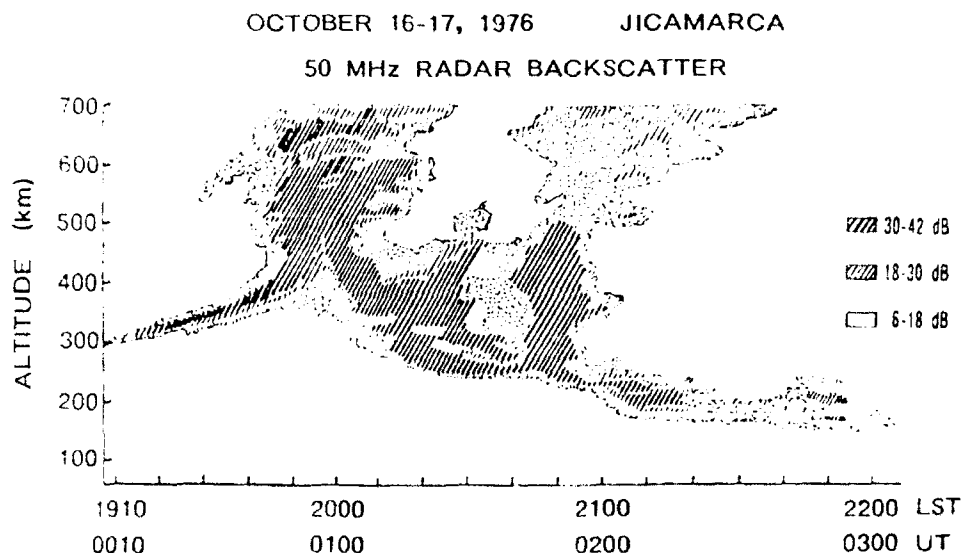


Figure 1. Radar backscatter from an extended (plume) region was observed in Jicamarca, Peru on October 16-17, 1976 with the data shown in Basu and Aarons, 1977. The irregularities start with thin layer bottom side returns in the 300 km region then develop into the extended region. Returns from the region above 700 are not noted because of the sensitivity of the radar. However in all probability in this case there are irregularities above 700 km.

taken by R. Tsunoda, M. Mendillo, and S. Basu in a multi-disciplined campaign in the Pacific in 1988. As shown in Figure 2, on August 8 the ALTAIR radar indicated that only a thin layer had developed at the time shown (0846 UT) with returns observed from 200 to 400 km. The relatively sensitive instrumentation at 137 MHz on Wake Island 11° from the magnetic equator showed no scintillation. The less sensitive scintillation measurements at the site near the equator (Kwajalein Island) at 250 MHz showed intense activity at the greater than 20 dB level. However when a plume with a great extent in height (Figure 4) was observed on the radar on August 7th at 1205 UT, both sites showed intense levels of scintillations.

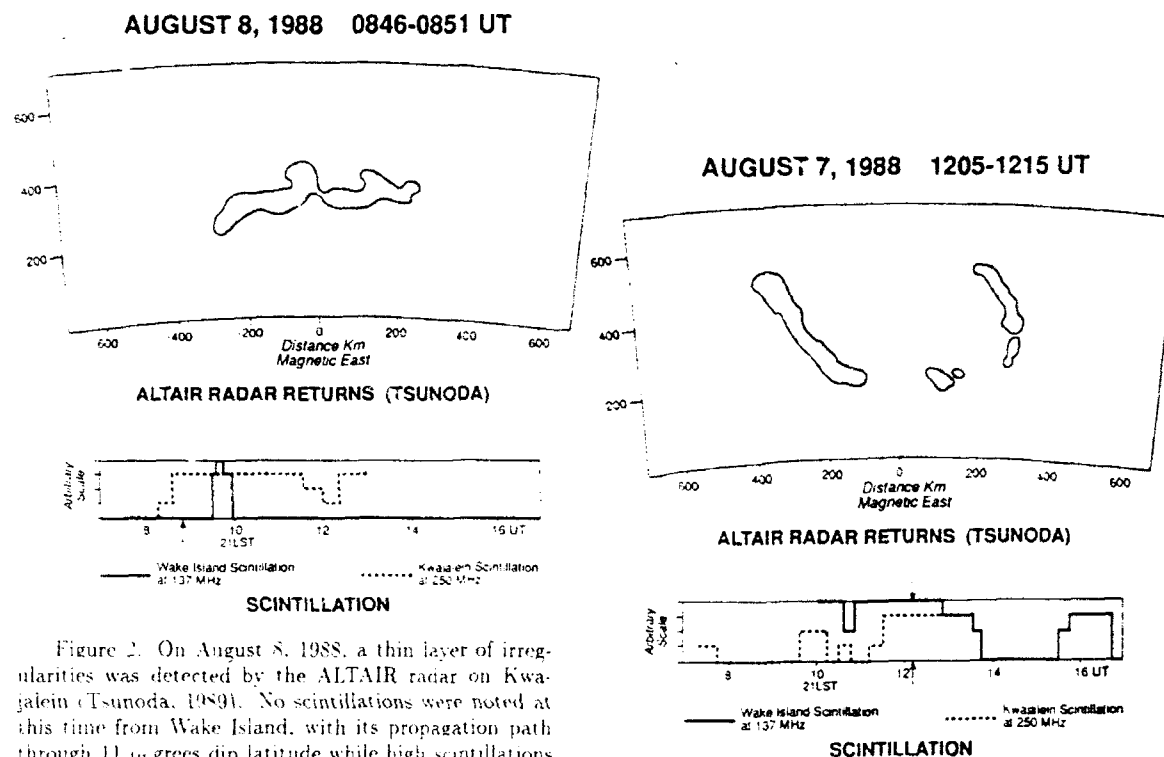


Figure 2. On August 8, 1988, a thin layer of irregularities was detected by the ALTAIR radar on Kwajalein (Tsunoda, 1989). No scintillations were noted at this time from Wake Island, with its propagation path through 11 degrees dip latitude while high scintillations levels were noted in the path from Kwajalein to another satellite. The higher latitude propagation path to the satellite at Wake Island missed the irregularity region while the path close to the magnetic equator at Kwajalein had high scintillation levels (S. Basu, and M. Mendillo, personal communication).

Figure 3. On August 7, 1988 through the irregularities extended in altitude as shown by the ALTAIR radar at the time period noted by the arrows, both propagation paths (Kwajalein and Wake) showed high levels of scintillation.

Figure 4 illustrates the concept that the disturbance on the magnetic equator follows the lines of force of the earth's field. The disturbance affects the electron density all along the magnetic field line. Thus a satellite at 900 km making in-situ measurements passing thru the plume would report observing irregularities; a satellite passing the longitude of the thin layer would not. While F-layer irregularities developed in both the thin and the extended layer cases shown, paths beyond 10° from the magnetic equator could not observe F-layer irregularities due to thin layer irregularities. A portion of the thin layers have been identified as Bottomside Sinusoidal structures by both ground and satellite observations.

#### D. SATELLITE MEASUREMENTS

There are two types of measurements made from satellites. In-situ observations measure parameters at the height of the measurements. Satellites with topside ionosondes however will sample lower altitude spread F as well as plumes extended in altitude. In either case the altitude of the satellite will affect the morphology developed by the satellite observations.

Although it appears from the radar data that the plumes extend only to altitudes of 700-900 km, satellites in the higher altitude region of 900 km and greater do measure irregularities at altitudes considerably higher than the radar sensitivity can observe; the plumes extend beyond the range of the radars.

The relatively thin layers were sampled in situ and reported by Valladares et al (1983) and Cragin et al (1985) in their observations using the AE-E satellite operating at 400-450 km. Thin layers extending 7000 km or more in the east-west direction were observed. These thin layers had been observed frequently in various campaigns (Basu and Aarons, 1977) but had not been studied in the manner that the plumes were studied. Certain thin layers, termed Bottomside Sinusoidal structures, have a dominant scale size of 1 km and appear predominantly in the summer solstice along the magnetic equator. They appear, according to the illustrations in Cragin et al. (1985), within plus and minus 5° of the magnetic equator.

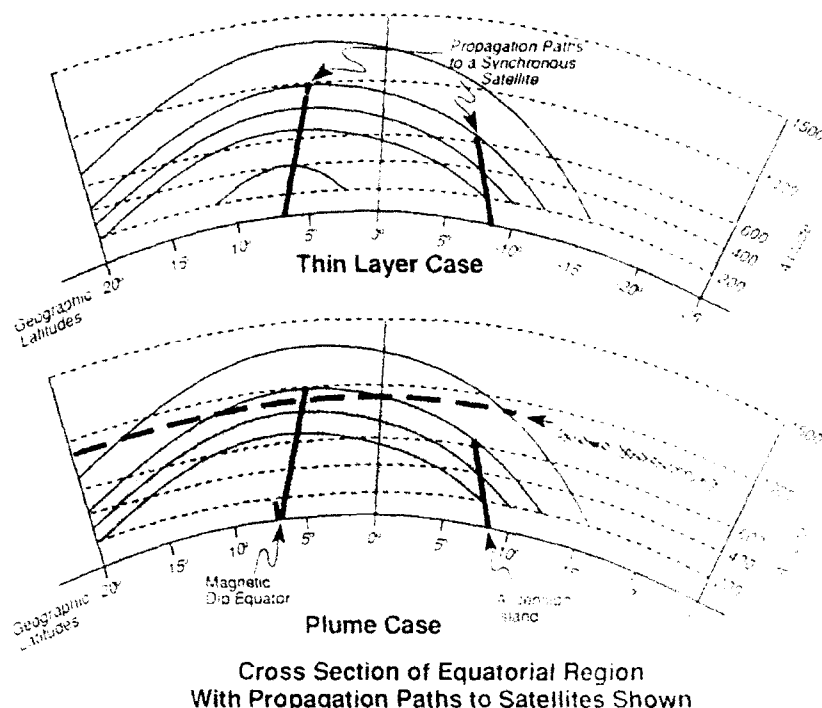
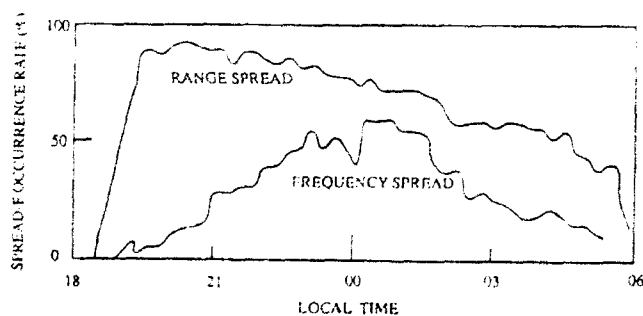


Figure 4. Propagation paths are shown through thin and extended layers of irregularities. The irregularities appear along the lines of force of the earth's field. A satellite at 900 km would only sample the irregularities over the magnetic equator. A lower altitude satellite might detect irregularities at both magnetic equator and anomaly latitudes.

#### E. SPREAD F

The discovery of spread F in the equatorial region was made with ionosonde observations. It was quickly found that there were two types of spread F, range and frequency. It is important to see which type was associated with other observations of F-layer irregularities such as scintillation of trans-ionospheric signals. Huang (1970) observing at 137 MHz near the anomaly region from Taiwan was able to state that "the correlation of scintillation with range spread is extremely high and strongly positive ( $r = 0.9$ ) while scintillation is only weakly correlated with frequency spread."

However that correlation is not valid near the magnetic equator. Dastguta et al (1983) observing scintillations close to the magnetic equator and spread F from Huancayo found that the continuous scintillation patches of a duration of 5-6 hours (thin layers), were associated with frequency spread. The occurrence patterns of range and frequency spread near the magnetic equator can best be shown in Figure 5 with Fortealeza, Brazil data and curves taken from Abdu et al, 1981. While this illustration is for a limited period, the data serve to show the predominance of range spread in the early part of the night and the increase in occurrence of the frequency spread in the post-midnight time period.



Fortealeza, Brazil, Nov.-Feb.

[Abdu et. al (1981)]

Figure 5. The occurrence of range and frequency spread is shown for one season for Fortealeza, Brazil (Abdu et al, 1981). Maximum occurrence takes place for this longitude in the post-sunset hours while the frequency spread maximum occurs in the midnight time period.

Thus we have two distinct types of F layer irregularities, a thin long lasting layer associated with frequency spread F and plumes associated with range spread F. In addition, in some radar plots, there is evidence of plumes not fully extended in altitude (Basu and Aarons, 1977).

## F. MORPHOLOGY FROM SCINTILLATION MEASUREMENTS

### F.1 Irregularity Intensity and Scintillation Measurements

For the measurements made by studying the fluctuations of a radio beacon (scintillation), several factors should be borne in mind. The scintillation technique of recording the fluctuations of amplitude and phase of signal from satellites is an integrated measurement with all the irregularities in the propagation path affecting the trans-ionospheric signal. The amplitude of fluctuation is inversely proportional to frequency to the power 1.5 to 2. Thus the lower frequencies are more sensitive than the microwave observations. The use of lower frequencies such as 137 MHz is a boon in cataloging weak irregularities compared to the less sensitive observations at 1.5 GHz. However there are technical problems in the data at both ends of the observing spectrum. The lower frequency is very sensitive compared to the 1.5 GHz signal but is limited in its excursions in the data sets to be described to 20-25 dB peak to peak fluctuations. This is due to saturation of its signal to noise range. The upper frequency however seldom shows excursions greater than 7 dB maximum near the equator. Much of the 1.5 GHz data base was done with extremely small excursions. In predominantly anomaly data taken at 4 GHz, fair used levels of .5 dB and greater. Other data in the literature have used excursions greater than 1.5 dB or 2 dB; this in data sets where fluctuations of less than 1 dB were considered noise of the system.

Measurements at 137 MHz do show some differences from those taken at 1.5 GHz. However, we shall see that monthly occurrence trends are not different in the available data for two separated frequencies with extensive data bases near the magnetic equator.

#### a. HUANCAYO (75° West)

Perhaps the largest body of data for statistical purposes consists of scintillation measurements taken at Huancayo, Peru. The data from this site, located 12° south of the geographic equator and on the magnetic equator with a declination of 4° to 8° West (depending on the longitude of the satellite under observation), has the advantage of comparison with radar and spread-F measurements taken at Jicamarca, Peru.

The scintillation data sets available range from 10 MHz (BEB and BEU satellites at 100 km altitude) to 1.54 GHz signals from MARISAT. Thus one can note the apparent and real differences in morphology of scintillation data as a function of frequency as well as month, solar flux, and magnetic index.

An extended period of observations was made using the AFS series of satellites transmitting at 137 MHz. The data from January 1968 through February 1974, with Kp from 0 to 2+, show the occurrence pattern to have maximum levels in February and in October-November with a high occurrence of 6 dB peak to peak fluctuations from October through April in sunspot maximum years (Figure 6a). A minimum level was found in June and July. In Figure 6b a similar pattern is seen for sunspot minimum years with a somewhat high occurrence January 15-March 15 than in September-October.

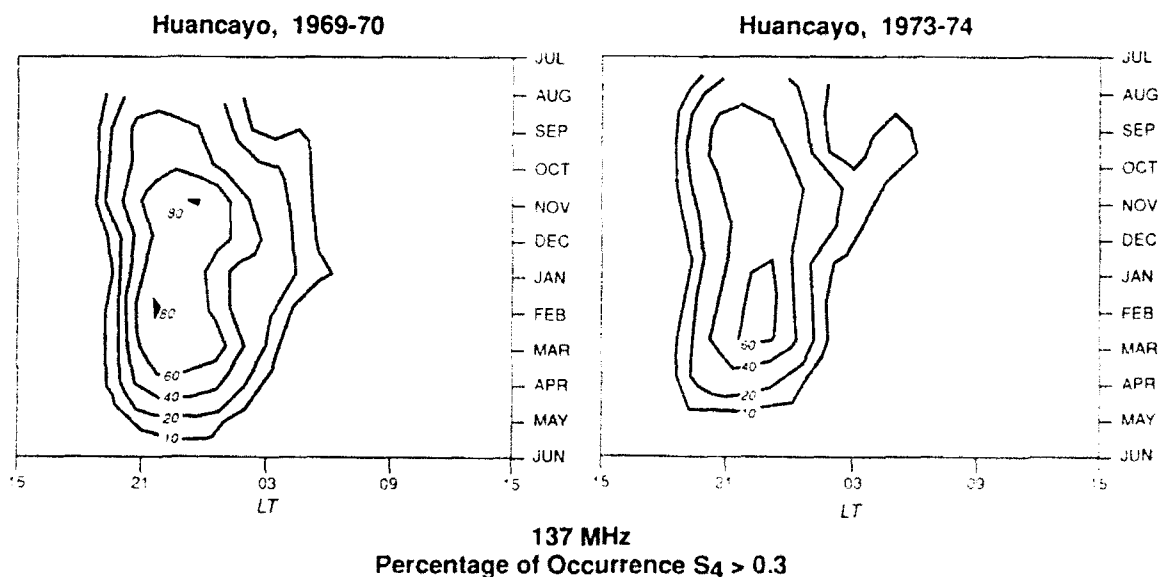


Figure 6. Occurrence levels from Huancayo, Peru. The occurrence patterns are marked with the mid months labels for each month noted. In tracing the pattern of scintillation occurrence during periods of low magnetic index at 137 MHz, the high solar flux years show a maximum between October and April. However with solar minimum data one can see the maximum occurrence January 15-March 15 with a slightly lower occurrence in September and October and with a deeper dip in level between November and January. June and July show extremely low levels of occurrence.



In Figure 7 taken at 1.54 MHz as shown in Basu et al. 1988, during a year of high solar flux, the pattern is quite similar to that taken at the lower frequency with equinox maxima, a shallow null in the November to January period, and extremely low to zero occurrence from the middle of May to the middle of August.

For this longitude, the morphology consists of peaks at the equinoxes, with the Feb-March peak higher than the September-October peak. A dip in the occurrence pattern takes place during the November-January period. Extremely low occurrence takes place from the middle of May to the middle of August.

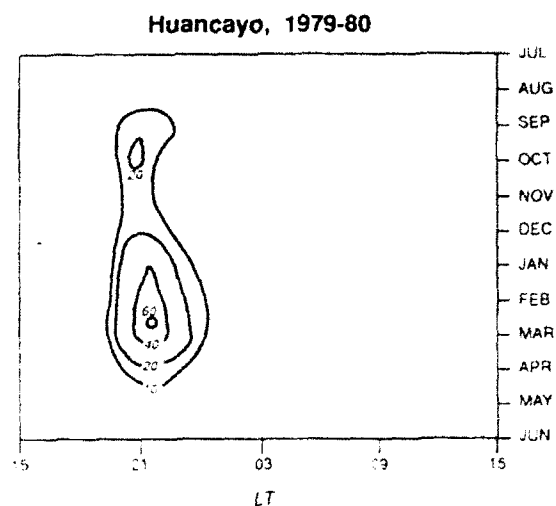


Figure 7. At 1.5 GHz the pattern of scintillation occurrence retains the same form as shown in the previous diagram.

1.5 GHz  
Percentage of Occurrence > 2dB

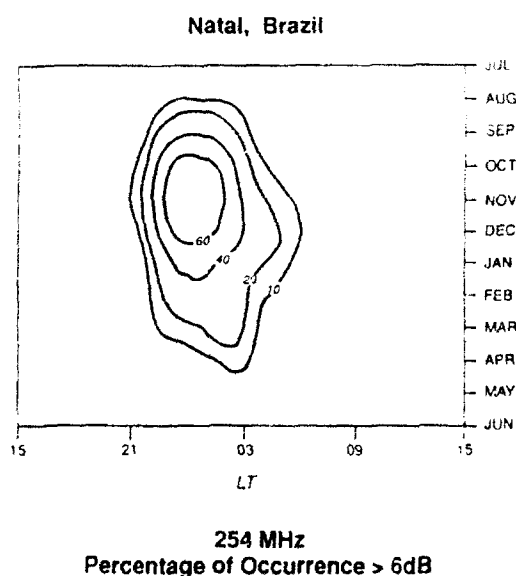
Basu, Su. et al., 1988

#### b. NATAL, BRAZIL

In Figure 8, Natal at 30° West, shows an occurrence pattern at 250 MHz with maximum from mid October to mid December. Lowest occurrence was noted in the May to August period. All magnetic indices were included in this set.

#### c. ACCRA, GHANA

The patterns of average scintillation level from Accra for 1967-1971 data at 137 MHz are shown by Koster (1972) in Figure 9. A slightly higher average level takes place from October to December (depending on year) than in February-March with a distinct low in June and July.



254 MHz  
Percentage of Occurrence > 6dB

Figure 8. At 250 MHz the pattern of occurrence for Natal, Brazil at 30° West shows a maximum in the October to January months.

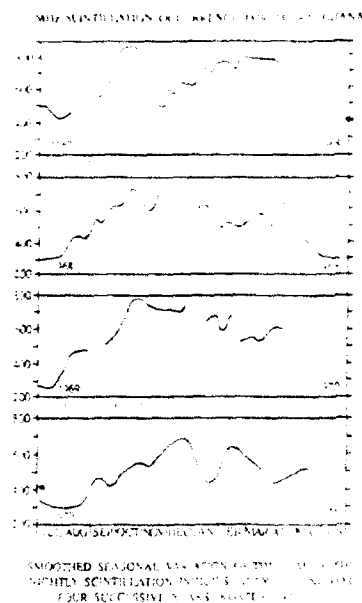


Figure 9. A long term series of 137 MHz measurements from Accra, Ghana near the Greenwich meridian yields the pattern shown in this figure i.e. a maximum from February to May, high levels between October and May, and the lowest levels from June to August (Koster, 1972).

#### d. INDIA

There is a limited amount of data in this longitude region, from within plus or minus 7°. Range spread data from Sastri and Murthy, 1975 as shown in Figure 10 indicates a minimum in the pre-midnight hours in June-August and in November and December. The data was taken in 1968, a year of relatively high solar flux.

#### e. MANILA

Data available for 1981, a year of high solar flux, show equinoctial maxima with a shallow minimum in June and July and a deep minimum in December and January during the year of high solar flux, 1981 (Figure 11a). In a year of low solar flux, 1986, the occurrence is considerably less but the equinox maxima hold (Figure 11b).

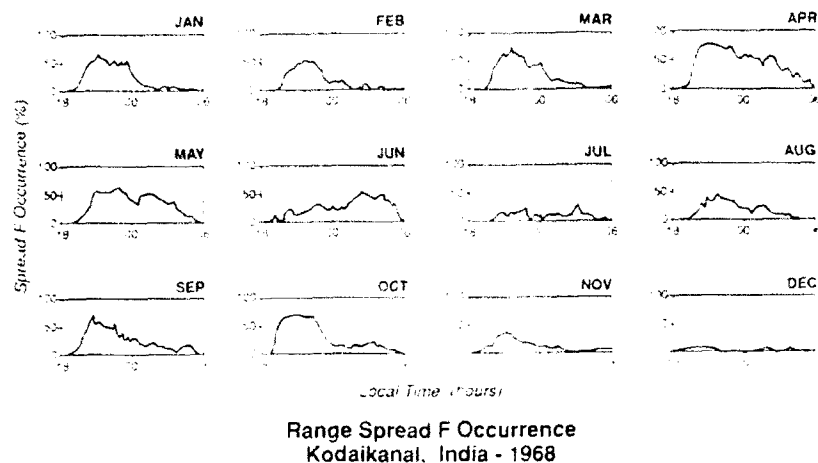


Figure 10. Long term scintillation data from observations close to the magnetic equator are difficult to come by in the Indian sector with some reported observations not extending to a year. However the occurrence of range spread-F is readily seen to have a minima in the post-sunset time period in June, July and August and in November and December (Sastri and Murthy, 1975). This essentially yields predominantly equinoctial maxima.

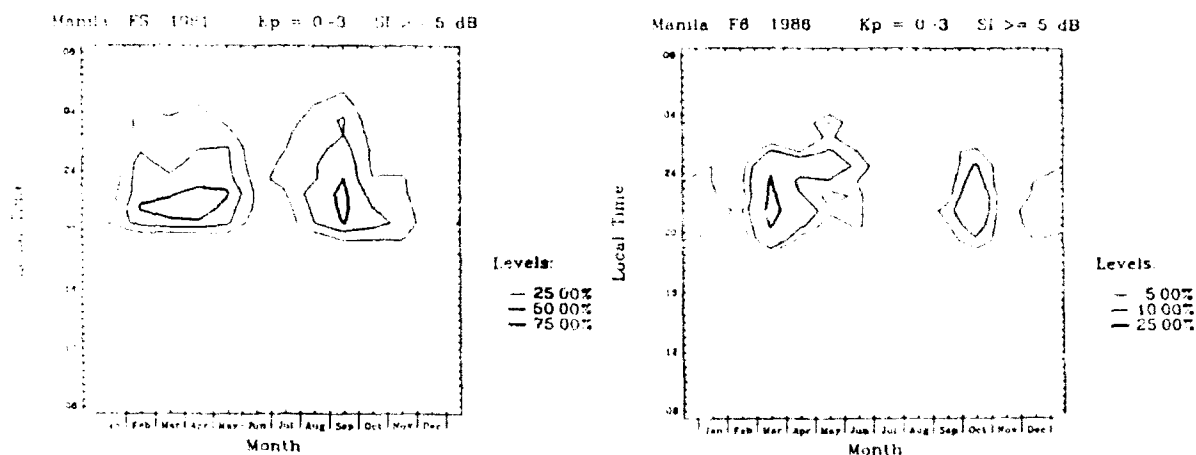


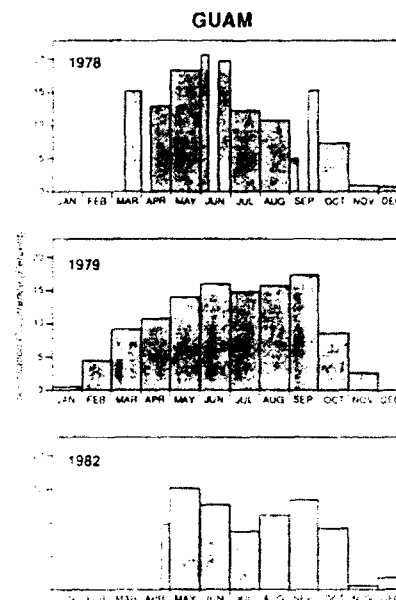
Figure 11. Using 250 MHz data taken over several years from Manila, it can be seen that the pattern is indicative of equinoctial maxima both in years of high sunspot number (1981) and in years of low sunspot number (1986). While the pattern is similar, it should be noted that the occurrence differs considerably between the two years.

#### f. GUAM

In a series of observations of two satellites both transmitting in the frequency range of 250 MHz, Paulson (1980) took measurements during years of high and low solar flux. The occurrence for 1978 and 1979 for the Pacific Ocean Satellite, the satellite where the propagation intersection is closely related to the longitude of Guam, is shown in Figure 12. Clearly the levels are high from April thru October. There is a shallow dip in June and July. A minimum of occurrence takes place from November through February. Additional data taken in 1982 and 1983 indicate a similar pattern.

Observations of a second satellite were taken for a propagation path farther to the west, the Indian Ocean Satellite. From the example of scintillations with this path, as shown in Paulson (1980), there may be, at times, tropospheric fluctuations at the low angle of observation. In addition the azimuthal angle was such that the propagation path was considerably to the west. Looking to the west with this series of observations of the Indian Ocean satellite, the pattern reveals a more equinox oriented structure.

Figure 12. From Guam over a period of several years, the occurrence of scintillations from a satellite near its meridian transmitting at 250 MHz was recorded by Paulson (1980). The pattern shows a clear maximum in the April to October period. It is likely that January to mid-April in 1982 was a period of no data.



#### g. KWAJALEIN

A set of data is available from Livingston (1980) for Kwajalein Island. The data taken with a high inclination 1900 km satellite in the midnight time period are shown in Figure 13. The occurrence levels for the two frequencies are quite similar in outline. Two aspects of this diagram bear notice. There is a dip in occurrence level in June and July compared to May and August. The level between October and February is particularly low. This diagram indicates that the thin layer BSS does not contribute significantly to the occurrence. At this longitude the summer maximum of BSS would increase the levels in the June solstice. Since the higher frequency shows the same pattern (as does E Band data which are shown in the Livingston paper), thin layers do not seem to change the form of this morphology.

#### Summary

- India and Manila show definite equinox maxima.
- Huancayo, Natal, and Accra show equinoctial maxima with high levels of occurrence in the November-December period and low levels of occurrence in the June-July period.
- Guam and Kwajalein show high levels somewhat later than the March equinox and earlier than the September equinox. There is a shallow dip in occurrence in June in several data sets but with high levels for July and August.
- There is contamination in some data from thin bottomsides irregularities but in general the scintillation data show similar trends from 137 MHz to 1.5 GHz.

#### G. THEORIES FOR THE LONGITUDINAL PATTERN OF THE IRREGULARITIES

While it is clear that in the study of geophysics, there is no perfect correlation of phenomena, it is also clear that the theory must be able to explain the broad outlines of the occurrence patterns of phenomena. The occurrence pattern of equatorial F-layer irregularities includes equinoctial maxima with very low occurrence in the solstices at some longitudes, equinoctial maxima with a relatively shallow dip in occurrence in the June-July periods at some longitudes such as Kwajalein and equinoctial maxima with a relatively shallow dip in occurrence in the November-December months in longitudes such as Huancayo. Two of the proposed hypotheses have been those from Tsunoda (1988) and from Matuyama and Matuura (1984).

#### COMMENTS: 1984 paper by Matuyama and Matuura

By observing with a topside ionosonde, Matuyama and Matuura (1980) were able to determine the morphology of the intensity of spread echoes presumably from all ranges below the 1100 km satellite. They measured strength of irregularities. From their data in the period, August 1978 to June 1979, with sunspot numbers ranging from moderate to high, they found a relatively high occurrence over the magnetic equator region and a lower occurrence over the anomaly region. The data set is shown in Figure 14.

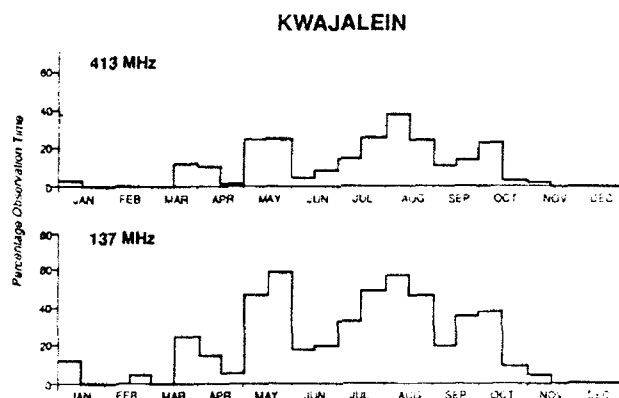


Figure 13. A one year series of observations of the Wideband satellite from Kwajalein in the time period around midnight yielded the occurrence pattern shown (Livingston, 1980). From this figure the general occurrence pattern is seen to be similar at 140 MHz and at 113 MHz. The occurrence for 1500 MHz shown in the Livingston paper showed similar trends.

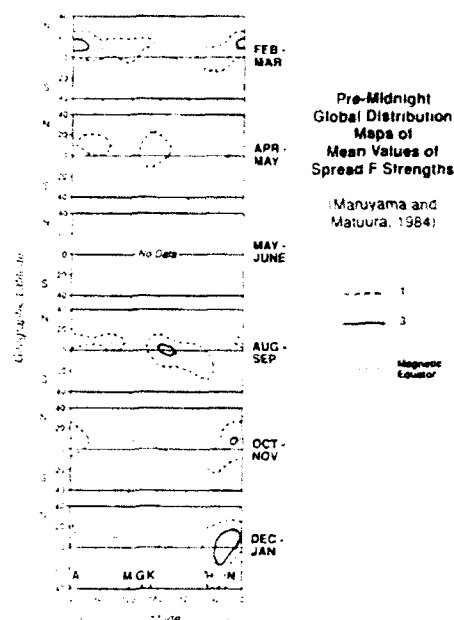


Figure 14. The data set from Maruyama and Matuura (1980) for the period before midnight is illustrated. There are gaps in the coverage but the general pattern of spread-F strengths matches those of spread-F occurrence. In the separation of months there seems to be an absence of emphasis on equinoctial high levels in all areas.

In Maruyama and Matuura (1981), the authors develop the concept that the declination of the observing station determines what occurrence pattern of F layer irregularities will be observed; declinations are given in Figure 15. Their conclusion as stated in their abstract was: "During the northern winter periods, maximum enhancement at the Atlantic longitudes of large westward geomagnetic declination and during the northern summer at the Pacific longitude of large eastward declination." The stress therefore is on the solstices.

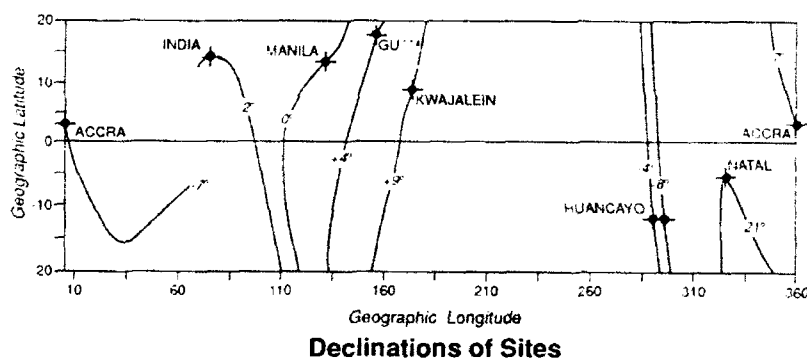


Figure 15. The sites of the various scintillation observational points and declinations from these sites are shown in this figure. It can be seen from this diagram and from the scintillation data that the westward declinations of Accra, Natal, and Huancayo tend to emphasize the September to March period while the eastward sites tend to emphasize the March to October period.

There are several problems with using the data set shown in the paper. Spread F strength is calculated but no differentiation is made between range and frequency spread. That is acceptable if only times with a few hours after sunset are used; range spread dominates in the immediate post-sunset time period. However the use of other times will describe all Spread-F. The seasonal pattern will be distorted relative to the seasonal pattern shown by scintillation or in-situ satellite measurements.

The main hypothesis however needs questioning. While we find similar patterns for similar declinations such as Huancayo and Accra and for Guam and Kwajalein, there is no comment on the equinoctial maxima at Huancayo and Accra. There is also a high occurrence particularly near the September equinox in the Central Pacific.

#### Comments on the Tsunoda (1988) Paper

Tsunoda (1988) on the other hand emphasizes equinox periods, moving close to or far from their ordinary dates. In the abstract he states "The seasonal maxima in scintillation activity coincide with the times of year when the solar terminator

(at E-layer heights) is most nearly aligned with the geomagnetic flux tubes. Hence the seasonal pattern of scintillation activity, at a given longitude, becomes a simple deterministic function of the magnetic declination and geographic latitude of the magnetic dip equator".

We have made a detailed examination of the Tsunoda data set. In Figure 16, we have plotted the sunset nodes that Tsunoda has calculated for each of the data sets relative to this study. The general structure is such that the nodes of Guam and Kwajalein are closer to the June solstice than the other nodes; the maximum of activity however is between these two points. The nodes of Natal, Accra, and Huancayo are closer to the December solstice and the maximum of activity for these sites is near the December solstice. The nodes for Natal are particularly close to the December solstice; Natal does have a somewhat higher level for this solstice than the other two sets of data. This compares favorably with the tendency of the occurrence pattern to move towards maxima in the various regions.

However the nodes for Huancayo and Guam are very close together but the patterns of scintillation activity for these two locations are quite different. In particular the high occurrence in December and January in Huancayo data is in direct contrast with the very low level of occurrence in the Guam data for those months. Similarly there is a contrast in the June/July data for the two sets. This holds true for the higher frequency data, suggesting that the thin layers have not made a significant contribution to the pattern.

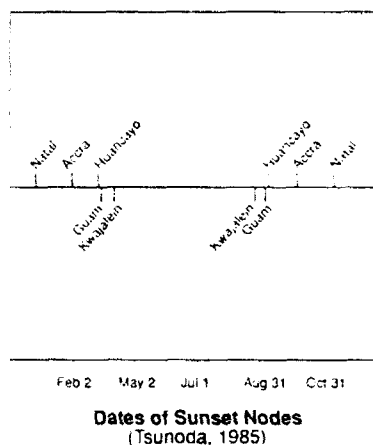


Figure 16. The sunset nodes as shown in Tsunoda (1988). Nodes indicate the time when E layer sunset for the data sites used was simultaneous for both ends of the field line running thru the site.

#### Questions to Be Answered by Theory

The stress of these two papers is either on the solstices or on the predominantly equinoctial maxima. However the data set demands that the questions to be answered by the developers of hypotheses for the physics of the longitudinal aspects of the morphology are:

- (1) Why do the equinox months have high levels of occurrence over all longitudes?
- (2) Why are there high levels of occurrence in the Central Pacific Sector in the July-August period and in the 0-75° West sector in the November-December period?
- (3) Why is there very low occurrence in November and December in the Central Pacific Sector and in July and August in the 0-75° West Sector?

#### CONCLUSIONS

A cartoon as to our version of the occurrence pattern of primarily extended F layer irregularities as a function of longitude is given in Figure 16. The evaluation of the data yields the occurrence for Huancayo, Ghana, India, and Guam-Kwajalein. A curve that is omitted for lack of data is that for the 120-140° West longitude region. The concept that might be developed is that there should be a region where the transition between two patterns with different solstitial maxima would dictate equinoctial maxima.

We have shown the plethora of measurements of F-layer irregularities in the equatorial region performed by various techniques. Irregularities that can be observed by radar may not be observed by the ionosondes. Irregularities that may be weak may not be observed by UHF scintillations but can be observed on VHF transmissions. Satellite scintillation techniques cannot always distinguish between thick and thin layers of irregularities; some of the thin layers yield very high scintillation indices at 250 MHz. Irregularities over the magnetic equator may not be observed by 6300 Å depletion techniques but can be observed readily in the anomaly region. The irregularities created can be strong or weak, of large or small scale, extended over a thin or thick layer, confined in a narrow east-west or extended region, old or new.

Each of the two suggested mechanisms embrace only a portion of the data set. A full concept should explain the physics behind the equinoctial and solstitial generation of irregularities. In addition there is also need to explain the differing mechanisms which provide thin or extended irregularity regions.

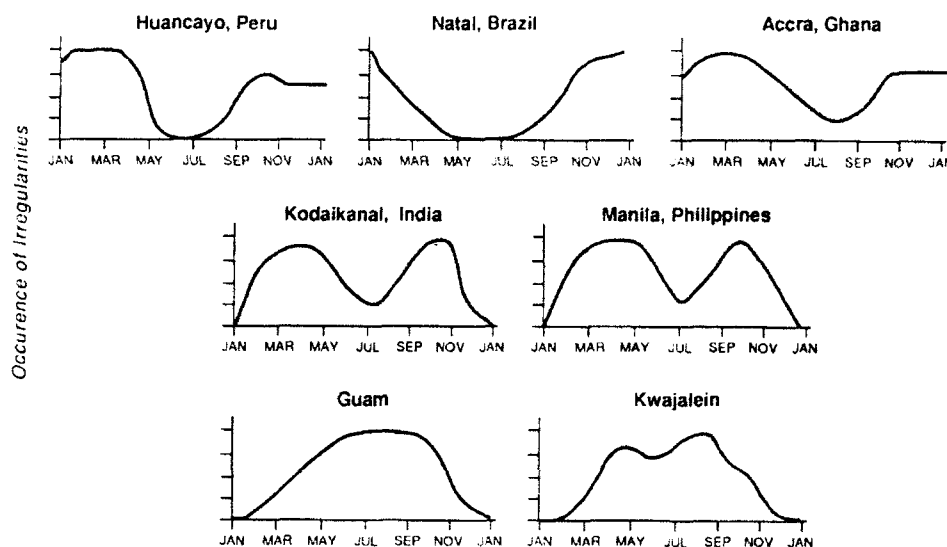


Figure 17. Using the scintillation data as indicating the F layer irregularities from all types of layers but predominantly from extended layers, it is possible to develop a sense of the morphology. We have attempted to do so in this figure. There are caveats however. The maximum amplitudes of the excursions shown are arbitrarily set as the same for all areas. In addition there is no distinction between high and low solar flux years. The pattern attempts to merge the two occurrence patterns from the data set but there may well be differences.

ACKNOWLEDGEMENTS: The studies of J. Aarons are supported by the Office of Naval Research.

#### BIBLIOGRAPHY

- Abdu, M.A., J.A. Bittencourt, and I.S. Batista, Magnetic declination control of equatorial F region dynamo electric field development and Spread F, *J. Geophys. Res.*, 86, 11442-11446, 1981.
- Basu, S. and J. Aarons, Equatorial irregularity campaigns Part 1: Correlated scintillation and radar backscatter measurements in October 1976, AFGL, Hanscom AFB, MA, AFGL-TR-0264 #617, 1977.
- Basu, Su., E. MacKenzie and S. Basu, Ionospheric constraints on VHF/UHF communication links during solar maximum and minimum periods, *Radio Science*, 23, 363-378, 1988.
- Cohen, R. and K.L. Bowles, On the nature of equatorial spread F, *J. Geophys. Res.* 66, 1081-1106, 1961
- Cragin, B.L., C.E. Valladares, W.B. Hanson and J.P. McClure, Bottomside sinusoidal irregularities in the equatorial F region, 2. Cross-correlation and spectral analysis, *J. Geophys. Res.*, 90, 1721-1734, 1985.
- DasGupta, A., J. Aarons, J.A. Klobuchar, Santimay Basu and A. Bushby, Ionospheric electron content depletions associated with amplitude scintillations in the equatorial region, *Geophys. Res. Lett.* 9, 147-150, 1982.
- Huang, Chun-Ming, F-Region irregularities that cause scintillations and spread-F at low latitude, *J. Geophys. Res.* 75, 4833-4841, 1970
- Koster, John R., Equatorial scintillation, *J. Planet. Sp. Sci.* 10, 1999-2014, 1972.
- Livingston, R.C., Comparison of multi-frequency equatorial scintillation, *Radio Sci.*, 15, 801, 1980.
- Maruyama, T. and N. Matuura, Global distribution of occurrence probability of spread echoes based on ISS-b observation, *J. Radio Res. Lab.*, 27, 201-216, 1980.
- Maruyama, T. and N. Matuura, Longitudinal variability of annual changes in activity of equatorial spread F and plasma bubbles, *J. Geophys. Res.*, 89, 10903-10912, 1984
- Paulson, M.R., Equatorial scintillation of satellite signals at UHF and L-Band for two different elevation angles, NOSC TR 543, 1980.
- Sastri, J.H. and B.S. Murthy, Spread-F at Kodaikanal Ann. *Geophys.* 31, 285, 1975
- Tsunoda, R.T., Control of the seasonal and longitudinal occurrence of equatorial scintillations by the longitudinal gradient in integrated E region Pedersen conductivity *J. Geophys. Res.* 90, 447-456, 1985
- Valladares, C.E., W.B. Hanson, J.P. McClure and B.L. Cragin, Bottomside sinusoidal irregularities in the equatorial F region, *J. Geophys. Res.*, 88, 8025, 1983

## DISCUSSION

### G. HAGN

The old vertical-incidence sounder (e.g., C-2/C-3/C-4) recorded spread F at many locations around the world. SRI (then Stanford Research Institute) operated such a sounder in Bangkok, Thailand in the early and mid-1960s (sunspot minimum in 1963 onward until 1967) and generated synoptic data on the probability of occurrence of both equatorial sporadic E and spread F. Could these tabulated statistics be of use to your study and supply another geographical sampling point, or is there a problem separating the range and frequency spreads?

### AUTHOR'S REPLY

Yes. There is a data gap between Manila and Indian longitudes. To be of use the data would have to be separated into range and frequency spread occurrence.

### R. BENSON

1. Please elaborate on the difference between low and high-latitude irregularities, i.e., that the low-latitude irregularities are much stronger.
2. How did you combine data from different techniques in your summary figure?

### AUTHOR'S REPLY

1. At auroral latitudes, intensity of scintillation signals at  $\sim 1$  GHz is rarely greater than a few dB. (At equatorial anomaly latitudes, there are frequently signal variations of greater than 10 dB on 4 GHz, a frequency that is less sensitive to scintillation). No in-situ data have either confirmed or rejected this concept.
2. Most of the data from radar is limited to periods of irregularity development and therefore not useful for morphology. The in-situ data published is also limited; I did use some. I used spread F, scintillation (basically), and scintillation on radar signals.

## IRREGULAR MEDIA EFFECTS ON RADIOWAVE SIGNALS USED IN NAVIGATION AND POSITIONING SYSTEMS

Ezekiel Bahar  
Electrical Engineering Department  
University of Nebraska-Lincoln  
Lincoln, NE 68588-0511 U.S.A.

### ABSTRACT

Radiowave signals, over a very broad range of frequencies, are used extensively in navigation and propagation systems. These signals that propagate to very large distances from the transmitter across the earth's surface or through the ionosphere are significantly affected by media irregularities. Some of these irregularities are fixed while others fluctuate due to, for instance, the diurnal variations in the ionosphere.

The magnitude and phase of CW signals are affected by the medium through which they propagate. Since these effects are frequency dependent, transient (pulsed) signals also undergo distortions and delays, which have a very significant impact on navigation and positioning systems. Thus when CW signals are used for navigation or positioning, it is necessary to predict the deviations in the phase of the received signals (phase anomalies) due to medium effects. When pulsed signals are used for navigation and positioning, it is necessary to predict the signal delays due to the medium effects. Moreover for given propagation paths, it is necessary to avoid the use of carrier frequencies that are more prone to signal distortions which make it very difficult to predict the time of arrival of the radio signals.

It is necessary to obtain accurate predictions of the received signals as the demands on the navigation system become more stringent. Thus for most relevant propagation problems, it is not sufficient to use idealized models of the propagation medium. The models that need to be considered consist of layers of nonuniform thickness with varying electromagnetic parameters. Since the ionosphere also needs to be considered, the medium is generally considered to be both inhomogeneous and anisotropic. For these irregular models of the curved earth-ionosphere propagation environment it is not possible to obtain standard separable solutions for the propagation problem. Furthermore since the solutions should be valid for a very broad range of radio frequencies, the physical/geometrical optics (high frequency) approach or the small perturbation (low frequency) approach cannot be used in general.

In this work, a full-wave approach is described in detail. Radiowave propagation over irregular propagation paths is examined in detail. The phase anomalies and pulse delays and distortions for radio signals transmitted across hills and valleys on the earth's surface or over the oceans and through irregular layered models of the earth's crust and the ionosphere are determined using the full-wave approach.

### 1. INTRODUCTION

In Section 2, we present the solution to the problem of VLF radiowave propagation in a perturbed spherical model of the earth ionosphere waveguide. The ionospheric perturbation considered is a gradual day/night transition in the effective height of the ionosphere. The generalized telegraphists' equations are solved for the forward (transmitted) and backward (reflected) wave amplitudes along a propagation path before, across and beyond the transition region. Through the use of auxiliary functions, the original problem with mixed

boundary conditions has been converted to a problem with unmixed boundary conditions. The field undulations and the resulting phase anomalies due to ionospheric perturbations can be obtained from the wave amplitudes along the propagation path. The results are summarized in Table I in terms of the waveguide scattering matrix whose elements are the  $n$ th mode reflection and transmission coefficients for an incident  $m$ th mode of unit amplitude.

In Section 3, a perturbation on the earth's surface is considered. The transient excitation is a LORAN C Pulse from a magnetic line source (vertically polarized waves). Coupling between all three components of the full wave field expansion of the fields (the radiation field, the lateral wave and the surface wave) is considered. The distortion of the received pulse are due to the singularities of the wave transfer functions. Special attention is given to determine the time of arrival (associated with the third zero crossing of the LORAN C Pulse) of the received signal.

In Section 4, full wave solutions are derived for the radiowaves propagating through an inhomogeneous anisotropic model of the ionosphere. The ionosphere is characterized by a scattering matrix whose elements are the like and cross-polarized transmission and reflection coefficients for incident horizontally and vertically polarized waves. From these results the Faraday rotation as well as the signal distortion delays can be obtained. Numerical and analytical solutions to the problem are compared.

### 2. TRANSMISSION AND REFLECTION SCATTERING COEFFICIENTS FOR THE MODES IN AN IRREGULAR SPHEROIDAL MODEL OF THE EARTH-IONOSPHERE WAVEGUIDE

Using a full wave approach, the electromagnetic fields are computed for observation points along the propagation path in an irregular spheroidal model of the earth-ionosphere waveguide (Bahar 1976). The perturbation considered here results in the variation of the effective height of the ionosphere. On employing complete expansions for the electromagnetic fields (based on the Watson transformation, Watson 1918, Wait 1962) and imposing the boundary conditions at the irregular surface, Maxwell's equations are transformed into sets of coupled ordinary differential equations for the forward and backward propagating wave amplitudes. The wave amplitudes, which are subject to mixed boundary conditions at the input and output terminal surface  $\theta = \theta^A$  and  $\theta = \theta^B$  of the waveguide (see Fig. 1) are expressed in terms of auxiliary functions with unmixed boundary conditions to obtain rigorous numerical solutions to the differential equations at arbitrary points along the propagation path ( $\theta < \theta^A$ ,  $\theta^A < \theta < \theta^B$  and  $\theta > \theta^B$ ). Thus, the analysis can be used to compute the perturbed fields inside and outside the irregular regions for arbitrary excitations. Since low-frequency radiowaves are used for navigation and positioning, special attention is given to the phase anomaly due to the ionospheric perturbations.

For the illustrative example presented here, the azimuthally independent effective height of the ionosphere is given by



(see Fig. 1)

$$h(\theta) = \begin{cases} h^A, & 0 < \theta < \theta^A \\ \frac{1}{2}(h^A + h^B + (h^A - h^B) \cos[\pi(\theta - \theta^A)/(\theta^B - \theta^A)]), & \theta^A \leq \theta \leq \theta^B \\ h^B, & \theta^B < \theta < \pi \end{cases} \quad (1)$$

where  $h(\theta) = r_{0,1}(\theta) - r_{1,2}$  and

$$\begin{aligned} h^A &= 6 \times 10^4 m, \quad h^B = 9 \times 10^4 m \\ \theta^A &= 89.8^\circ, \quad \theta^B = 90.2^\circ \end{aligned} \quad (2)$$

The permittivity  $\epsilon$  and permeability  $\mu$  are given by

$$\epsilon = \begin{cases} \epsilon_0 = \epsilon_f(1 - i) \\ \epsilon_1 = \epsilon_f \\ \epsilon_2 = \epsilon_f(10 - i10) \end{cases} \quad (3)$$

$$\mu_0 = \mu_1 = \mu_2 = \mu_f \quad (4)$$

in which  $\epsilon_f$  and  $\mu_f$  are the permittivity and permeability for free space. The mean radius of the earth's surface is  $r_{1,2} = 64 \times 10^6$  m and the frequency is  $f = 15$  kHz (corresponding to the free space wavelength  $\lambda = 20$  km). Steady state  $\exp(i\omega t)$  time excitations are assumed.

The characteristic equation for the waveguide modes is

$$R^{Dr} R^{Ur} = 1 \quad (5)$$

where

$$R^{Dr} = R^D h_v^{(1)}(k_1 r_{1,2}) / h_v^{(2)}(k_1 r_{1,2}) \quad (6)$$

$$R^{Ur} = R^U h_v^{(2)}(k_1 r_{0,1}) / h_v^{(1)}(k_1 r_{0,1}) \quad (7)$$

in which  $R^D$  and  $R^U$  are the reflection coefficients at the earth-air interface and at the effective ionosphere boundary (Waite 1962, Bahar 1967). Precise numerical methods have been developed to solve the modal equations for the orders  $\nu$  of the spherical Hankel functions of the first and second kind  $h_v^{(1)}$  and  $h_v^{(2)}$  (Bahar 1980, Bahar and Fitzwater 1981). They are related to the mode propagation coefficients  $\beta_k \approx \nu + \frac{1}{2}$ . In Fig. 2, the loci of the first four (vertically polarized) TM complex mode numbers  $\nu_1, \nu_2, \nu_3$  and  $\nu_4$  are plotted in the complex plane with the effective height  $h(\theta)$  as the variable parameter from  $h^A$  to  $h^B$ . The dominant (earth detached mode)  $\nu_1$  is least attenuated at the output port  $B$ . (See Fig. 1.)

Analytical and numerical procedures used to obtain the solutions to the modal equation (5) are described in detail in the technical literature (Bahar and Fitzwater 1981).

For the sourceless perturbed region  $\theta^A < \theta < \theta^B$ , where  $dh/d\theta \neq 0$ , the coupled wave equations may be written in matrix form as

$$-\frac{d}{d\theta} \begin{pmatrix} a \\ b \end{pmatrix} = i \begin{pmatrix} \beta & \\ & -\beta \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} S^{BA} & S^{BB} \\ S^{AA} & S^{AB} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \quad (8)$$

where  $a$  and  $b$  are  $n \times 1$  matrices whose elements are the  $n$  forward and backward wave amplitudes  $a(\nu, \theta)$  and  $b(\nu, \theta)$ ,  $\beta$  is an  $n \times n$  diagonal matrix whose elements are the mode propagation coefficients  $\beta_k$  and  $S^{PQ}$  are  $n \times n$  scattering matrices whose elements  $S_{\nu\mu}^{PQ}$  are the differential transmission ( $P \neq Q$ ) and reflection ( $P = Q$ ) coefficients. For the  $m$ th

mode of unit amplitude incident at port  $A$  the mixed boundary conditions at  $\theta = \theta^A$  and  $\theta = \theta^B$  are

$$\begin{aligned} a(\theta^A) &= \begin{pmatrix} \delta_{1m} \\ \delta_{2m} \\ \dots \\ \delta_{nm} \end{pmatrix}, \\ b(\theta^B) &= \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \end{pmatrix} \end{aligned} \quad (9)$$

in which  $\delta_{nm}$  is the Kronecker delta. Furthermore,

$$\begin{aligned} a(\theta^B) &= \begin{pmatrix} T_{1m}^{BA} \\ T_{2m}^{BA} \\ \dots \\ T_{nm}^{BA} \end{pmatrix}, \\ b(\theta^A) &= \begin{pmatrix} R_{1m}^{AA} \\ R_{2m}^{AA} \\ \dots \\ R_{nm}^{AA} \end{pmatrix} \end{aligned} \quad (10)$$

where  $T_{km}^{BA}$  and  $R_{km}^{AA}$  are the waveguide junction transmission and reflection coefficients. To account for all possible excitations at port  $A$ , (8) is written as

$$-\frac{d}{d\theta} \begin{pmatrix} A \\ B \end{pmatrix} = i \begin{pmatrix} \beta & \\ & -\beta \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} + \begin{pmatrix} S^{BA} & S^{BB} \\ S^{AA} & S^{AB} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} \quad (11)$$

in which  $a_{km}$ , the forward scattered  $k$ th mode excited by an incident  $m$ th mode, and  $b_{km}$ , the backward scattered  $k$ th mode excited by an incident  $m$ th mode, are the elements of the  $n \times n$  matrices  $A$  and  $B$ , respectively. The differential equations (11) must be solved subject to the mixed boundary conditions,

$$A(\theta^A) = I, \quad B(\theta^B) = 0 \quad (12)$$

with

$$A(\theta^B) = T^{BA}, \quad B(\theta^A) = R^{AA} \quad (13)$$

in which the elements of the transmission and reflection matrices  $T^{BA}$  and  $R^{AA}$  are the  $k$ th mode transmission and reflection coefficients  $T_{kl}^{BA}$  and  $R_{kl}^{AA}$  for the  $l$ th mode of unit amplitude incident at port  $A$ . To facilitate the solution of (11) with the mixed boundary condition (12), define

$$\begin{pmatrix} A \\ B \end{pmatrix} \equiv \begin{pmatrix} F \\ G \end{pmatrix} T^{BA} \quad (14)$$

in which  $F$  and  $G$  are  $n \times n$  matrices. Since  $T$  is a constant matrix (11) reduces to

$$-\frac{d}{d\theta} \begin{pmatrix} F \\ G \end{pmatrix} = i \begin{pmatrix} \beta & \\ & -\beta \end{pmatrix} \begin{pmatrix} F \\ G \end{pmatrix} + \begin{pmatrix} S^{BA} & S^{BB} \\ S^{AA} & S^{AB} \end{pmatrix} \begin{pmatrix} F \\ G \end{pmatrix} \quad (15)$$

with the unmixed boundary condition

$$F(\theta^B) = I, \quad G(\theta^B) = 0 \quad (16)$$

The unknown waveguide junction transmission and reflection coefficients are given by

$$T^{BA} = F(\theta^A)^{-1}, \quad R^{AA} = G(\theta^A)T^{BA} \quad (17)$$

Thus by using the Runge-Kutta method (Abramowitz and Stegun 1964), rigorous numerical solutions may be obtained for the transmission and reflection coefficients by solving (15) subject to the unmixed boundary conditions at  $\theta = \theta^B$  (16) instead of solving (11) subject to the mixed boundary conditions. When port  $B$  ( $\theta = \theta^B$ ) is excited instead of port  $A$ , the above analysis is repeated on interchanging superscripts  $A$  and  $B$  to obtain the matrices  $T^{AB}$  and  $R^{BB}$  whose elements are the transmission and reflection coefficients  $T_{kl}^{AB}$  and  $R_{kl}^{BB}$ .

It has been shown that the scattering coefficients  $S_{kl}^{PQ}$  satisfy the exact reciprocity relationships for spheroidal waveguides (Bahar 1976). In view of the normalization used in this work, the reciprocity relations for the perturbed waveguide junction are

$$T_{kl}^{AB} = T_{lk}^{BA} \quad (18)$$

and

$$R_{kl}^{AA} = R_{lk}^{AA}, \quad R_{kl}^{BB} = R_{lk}^{BB} \quad (19)$$

The analysis presented in this section can be used to compute the electromagnetic fields in the perturbed regions  $\theta^A \leq \theta \leq \theta^B$  as well as in the unperturbed regions  $\theta \leq \theta^A$  and  $\theta \geq \theta^B$ . For instance, when port  $A$  is excited by mode  $m$  of unit magnitude,

$$H_{zm}(r, \theta) = \sum_n \frac{\psi_n(r)}{B_n} [a_{nm}(\theta) - b_{nm}(\theta)] \quad (20)$$

and

$$E_{rm}(r, \theta) = \sum_n \frac{Z_n(r)\psi_n(r)}{B_n} [a_{nm}(\theta) + b_{nm}(\theta)] \quad (21)$$

where  $a_{nm}$  and  $b_{nm}$  are the element of  $A$  and  $B$  defined in (11) and  $\psi_n(r)$  is the  $n$ th mode basis function,  $Z_n$  the mode impedance and  $B_n$  is the normalization coefficient (Bahar 1980).

For  $\theta \leq \theta^A$

$$a_{nm} = \delta_{nm} \exp[-i(\nu_m^A + 1/2)(\theta - \theta^A)] \quad (22)$$

$$b_{nm} = R_{nm}^{AA} \exp[i(\nu_m^A + 1/2)(\theta - \theta^A)] \quad (23)$$

and for the region  $\theta \geq \theta^B$ ,

$$a_{nm} = T_{nm}^{BA} \exp[-i(\nu_m^B + 1/2)(\theta - \theta^B)] \quad (24)$$

$$b_{nm} = 0 \quad (25)$$

in which  $\nu_m^A$  and  $\nu_m^B$  are  $m$ th the mode parameters at ports  $A$  and  $B$ , respectively.

The stability of the phase of the signals is important for navigational purposes, the phase anomalies are defined by

$$\Delta\phi_m^H = \arg \left\{ H(r, \theta) \exp \int_{\theta^A}^{\theta} i(\nu_m + 1/2)d\theta \right\} \quad (26)$$

$$\Delta\phi_m^E = \arg \left\{ E_r(r, \theta) \exp \int_{\theta^A}^{\theta} i(\nu_m + 1/2)d\theta \right\} \quad (27)$$

Thus  $\Delta\phi_m = \text{const}$  when no mode scattering occurs. The fluctuations in  $\Delta\phi_m$  are largest when two interfering modes are of approximately equal magnitudes.

The fields are least affected by the perturbation for  $m = 1$ . This is because the earth-detached mode tends to be guided along the air-ionosphere boundary, while the higher-order modes are reflected at the ionosphere and the earth boundaries. Furthermore, since the earth-detached mode is least attenuated, its magnitude is larger than the magnitudes of the scattered higher-order modes for all  $\theta$  when the incident mode is  $m = 1$  (Bahar 1980).

While it is difficult to excite the earth-detached mode from a ground-based antenna, the advantages of exciting this mode by elevated sources are very significant. The transmission and reflection scattering coefficients that characterize mode scattering by a perturbed ionosphere excited in both directions are presented in Table I. Since modes  $n \geq 5$  attenuate strongly over the propagation path, only four modes are considered in the analysis. The results are in very good agreement with reciprocity. These results can be used to determine the electromagnetic fields for arbitrary excitations.

### 3. TRANSIENT ELECTROMAGNETIC FIELDS TRANSMITTED ACROSS IRREGULAR MODELS OF THE EARTH'S SURFACE

Propagation of radiowaves in uniform layered models of the earth has been analyzed extensively in the technical literature. For horizontally stratified models, the electromagnetic fields can be expressed completely, in separable form, in terms of a continuous spectrum of radiation fields and lateral waves as well as a discrete set of surface waves (Bahar 1974).

When it is necessary to consider more realistic, irregular models of the propagation medium, consisting of inhomogeneous layers of nonuniform thickness, the electromagnetic fields cannot be expressed in separable form. In these cases, using the full wave approach, it is shown that incident plane waves are diffusely scattered in nonspecular directions and partially converted into lateral and surface waves.

The results of a steady state, full wave analysis of vertically polarized waves incident upon one-dimensionally rough surfaces (Bahar 1977, 1978b) are the basis for the investigation of transient electromagnetic response from irregular models of the earth's surface. In this work, the excitation is assumed to be a LORAN C pulse and the Fast Fourier Transform technique (Brigham 1974) is used to obtain the transient response.

Since the transfer function for dissipative media is not analytic, it is found convenient to express the total signal response as the sum of the contribution from the poles of the excitation transform and the contribution from all the singularities (poles and branch cuts) of the transfer function. Since the third zero crossings of the response to the LORAN C excitations are regarded as the effective arrival time of the pulse, it is not only necessary to determine the propagation delays but also the phase anomalies due to the irregular ground effects. Both the instantaneous response and the envelope of the response are obtained.

For a magnetic line source of intensity  $K$  at a large distance from the nonuniform boundary, the magnetic scattered radiation field is given by (Bahar 1978b)

$$H_z^s(x, y) = |H_{zp}| \exp[-ik_0(\rho_0 + \rho)](-i\omega/\omega_c) F(C^I, C^u) \cdot I(C^I, C^u, h, L) \quad (28)$$

where  $|H_{zp}|$  is the magnitude of the vertically polarized specularly reflected field for a flat, perfectly conducting surface at the carrier frequency  $\omega_c$ .

$$|H_{zp}| = |H_0^I| 2k_0 \cdot L C_0^u / (2k_0 \rho)^{1/2} \quad (29)$$

The magnitude of the vertically polarized incident wave at

the origin is  $|H_0^i|$ ;  $\rho_0 = (x_0^2 + y_0^2)^{1/2}$ ,  $\rho = (x^2 + y^2)^{1/2}$  are the distances from the source and field point to the origin (see Fig. 3) and  $k_0 = \omega_c(\mu_0\epsilon_0)^{1/2}$  is the free-space wave number at the carrier frequency  $\omega_c$ .

$$F(C^f, C^i) = \frac{2C_0^f(C_1^f C_1^i - S_0^f S_0^i)(1 - 1/\epsilon_r)}{[C_0^f + C_1^f/n][C_0^i + C_1^i/n][C_0^f + C_0^i]} \quad (30)$$

$$I(C^f, C^i, h, L) = \frac{1}{2L} \int_{-L}^L \exp[i(C_0^f + C_0^i)k_0 h + i(S_0^f - S_0^i)k_0 x] dx \quad (31)$$

where  $C_0^f$  and  $S_0^f$  are the cosine and sine of the angle of incidence,  $\theta_0^f$ , in the medium  $y > h$  and  $C_0^i$  and  $S_0^i$  are the cosine and sine of the scattering angle  $\theta_0^i$ . The corresponding expressions for the medium  $y < h$  are denoted by symbols with subscript 1 instead of 0. Thus,

$$u_{0,1}^f = k_{0,1} C_{0,1}^f = k_{0,1} \cos \theta_{0,1}^f \quad (32)$$

$$\beta_{0,1}^f = k_{0,1} S_{0,1}^f = k_{0,1} \sin \theta_{0,1}^f \quad (33)$$

in which  $k_{0,1} = \omega_c(\mu_{0,1}\epsilon_{0,1})^{1/2}$  and in view of Snell's law,

$$\beta_0 = \beta_1 = (k_{0,1}^2 - u_{0,1}^2)^{1/2}, \quad \text{Im}(\beta_1) \leq 0 \quad (34)$$

The intrinsic impedance is  $\eta_{0,1} = (\mu_{0,1}/\epsilon_{0,1})^{1/2}$ ,  $n$  is the refractive index, and the nonuniform boundary is given by

$$y = h(x), \quad -L \leq x \leq L \quad (35)$$

It is assumed here that the surface is gently undulating and slopes are small. In (28), the scattering phenomenon is represented by  $F(C^f, C^i)$ , which is a function of the angle of incidence  $\theta_0^f$ , the scatter angle  $\theta_0^i$ , and the ground relative permittivity  $\epsilon_r$ , and  $I(C^f, C^i, h, L)$ , which is the only term that depends on the expression for the rough surface  $h(x)$ . For a flat, perfectly conducting surface, and with  $\theta_0^f = \theta_0^i$ ,  $F(C^f, C^i) = 1$  and  $I(C^f, C^i, h, L) = 1$ . When medium 1 ( $y < h$ ) is highly conducting and the fields in the region  $y < h$  are of no practical interest, it is convenient to analyze the problem by characterizing the interface  $y = h(x)$  by an approximate surface impedance for waves of grazing incidence (Bahar 1972):

$$Z_s \rightarrow \eta_0(\epsilon_r - 1)^{1/2}/\epsilon_r = \eta_0 z_s \quad (36)$$

and  $F(C^f, C^i)$  in (30) is replaced by

$$F(C^f, C^i) = \frac{2C_0^f(1 - S_0^f S_0^i - z_s^2)}{(C_0^f + z_s)(C_0^i + z_s)(C_0^f + C_0^i)} \quad (37)$$

The full wave solutions for the scattered radiation fields satisfy the reciprocity relationships in electromagnetic theory. The scattered radiation fields vanish near the boundary ( $\theta_0^f \rightarrow \pi/2$ ). This agrees with the empirically well-known result that for any surface, regardless of roughness, a scattered radiation field will not exist at the surface except for grazing incidence (Heckmann and Spizzichino 1963). However, due to surface irregularities, the incident radiation fields excited by line sources at large distances are coupled into guided surface waves and lateral waves (see Fig. 3). The scattered surface waves and lateral waves could therefore be the predominant contributions to the total field near irregular boundaries.

The scattered surface wave due to line sources at large distances from the boundary is (Bahar 1977):

$$H_{xz}^s(x, y) = |H_0^i| \exp(-ik_0 \rho_0) (\dot{\omega}/\omega_c)^{1/2} i k_0 L F(C_s, C^i) \cdot I(C_s, C^i, h, L) \cdot \exp[-ik_0(S_0 s_x + C_0 s_y)] \quad (38)$$

where

$$F(C_s, C^i) = (C_1^i C_{1s} - S_0^i S_{0s}) / (C_0^i + C_1^i/n) \cdot n(1 + 1/\epsilon_r) \quad (39)$$

$$I(C_s, C^i, h, L) = \frac{1}{2L} \int_{-L}^L \exp[ik_0\{C_0^i h + (S_0 s_x - S_0^i x)\}] dx \quad (40)$$

in which the surface wave parameters are

$$S_{0s} = C_{1s} = [\epsilon_r/(1 + \epsilon_r)]^{1/2}, \quad C_{0s} = -(1 - \epsilon_r)^{1/2} \quad (41)$$

This solution for the scattered surface wave is also consistent with reciprocity relationships. Thus (38)-(41) also represent the scattered radiation fields due to an incident surface wave. For highly conducting medium 1, the surface impedance concept (36) can be used to characterize the boundary  $y = h(x)$  to obtain approximate expressions for the scattered surface wave. In this case, the function  $F(C_s, C^i)$  in (39) is replaced by

$$F(C_s, C^i) = z_s S_{0s}(S_{0s} - S_0^i) / (C_0^i + z_s)(1 - z_s^2)^{1/2} \quad (42)$$

in (38), the input term  $|H_0^i| \exp[-ik_0 \rho_0]$  is the incident radiation field and the output term  $\exp[-ik_0(S_0 s_x + C_0 s_y)]$  is the scattered surface wave (see Fig. 3). As in (28), the coupling phenomenon is represented by  $F(C_s, C^i)$  and  $I(C_s, C^i, h, L)$ .

The scattered lateral wave due to line sources at large distances from the boundary is (Bahar 1977):

$$H_{xz}^l(x, y) = |H_0^i| \exp[-ik_0 \rho_0] [(2L/\lambda_c)/2\pi(L_1/\lambda_c)^{3/2}] \cdot F(C_s^l, C^i) \cdot I(C_s^l, C^i, h, L) \cdot \exp[-i(k_1 L_1 + k_0 L_s)] \quad (43)$$

where

$$F(C_s^l, C^i) = (nC_1^i C_0^s + S_0^i S_0^s) / (C_0^i + C_1^i/n) n^{3/2} \epsilon_r \quad (44)$$

$$I(C_s^l, C^i, h, L) = \frac{1}{2L} \int_{-L}^L \exp[ik_0\{C_0^i h + (S_0^s - S_0^i)x\}] dx \quad (45)$$

in which  $\lambda_c$  is the free-space wavelength at the carrier frequency  $\omega_c$ .

$$\begin{aligned} \beta^s &= k_0 S_0^s = k_1 S_1^s = k_1 = k_0 n, \\ u_0^s &= k_0 C_0^s = k_0 \sqrt{1 - n^2}, \\ u_1^s &= k_1 C_1^s = 0 \end{aligned} \quad (46)$$

The lateral waves are of practical significance only for low-loss medium 1 where

$$\begin{aligned} n &= n' - in'' = S_0^s = \sin \delta = \sin(\delta' - i\delta'') \\ &\approx \sin \delta' - i\delta'' \cos \delta' \end{aligned} \quad (47)$$

Thus,

$$u_0^s y + \beta^s x = k_0 L_s + k_1 L_1 \quad (48)$$

and the distance  $L_s$  and  $L_1$  are (Fig. 3)

$$L_s = y/\cos \delta', \quad \text{and} \quad L_1 = x - y \tan \delta' \quad (49)$$

where  $\delta'$  is the critical angle for total internal reflection. In (43)  $|H_0^i| \exp[-ik_0 \rho_0]$  represents the incident radiation field

and the output term  $\exp[-i(k_1 L_1 + k_0 L_2)]$  is the scattered lateral wave (see Fig. 3). The coupling between the radiation field and the lateral wave is represented by  $F(C^\delta, C^s)$  and  $I(C^\delta, C^s, h, L)$ . The solution for the scattered lateral wave is consistent with reciprocity relationships, thus the expression (43) also represents the scattered radiation fields due to an incident lateral wave. The lateral wave contribution cannot be derived if the approximate surface impedance concept (36) is used. When the line source is near the boundary  $h(x)$ , it excites surface waves and lateral waves in addition to the radiation term. In this case, it is also necessary to consider the coupling between the lateral waves and the surface wave at the irregular boundary. The surface wave generated at the irregular boundary  $h(x)$  by an incident lateral wave is

$$H_{il}^s(x, y) = [K i \omega \epsilon_0 2\pi / (-2\pi i k_1 x_0)^{3/2}] \cdot \exp[-ik_0(C_0^\delta y_0 - S_0^\delta x_0)] \cdot F(C_s, C^\delta) I(C_s, C^\delta, h, L) \cdot \exp[-ik(C_0 s y + S_0 s x)] \quad (50)$$

where

$$F(C_s, C^\delta) = [C_0^\delta + 1] \{ \epsilon_r / (1 - \epsilon_r^2) \}^{3/2} \quad (51)$$

$$I(C_s, C^\delta, h, L) = \frac{1}{2L} \int_{-L}^L \exp[-ik_0(C_0^\delta h - (S_0^\delta - S_0 s)x)] dx \quad (52)$$

and for  $-x_0 \gg L$  the amplitude of the unperturbed lateral wave at the origin is

$$H_{il}(0, 0) = \frac{-K i \omega \epsilon_0}{(1 - \epsilon_r) 2(2\pi)^{1/2} [-ik_1 x_0]^{3/2}} \cdot \exp[-ik_0(C_0^\delta y_0 - S_0^\delta x_0)] \quad (53)$$

Hence in (50)  $\exp[-ik_0(C_0^\delta y_0 - S_0^\delta x_0)]/[k_1 x_0]^{3/2}$  corresponds to the incident lateral wave while the output term  $\exp[-ik_0(C_0 s y + S_0 s x)]$  corresponds to the scattered surface wave.

This solution for the coupling between the surface wave and the incident lateral wave by irregular surfaces is also consistent with reciprocity relationships in electromagnetic theory. Thus, (50) also represents the lateral wave generated at the irregular boundary by an incident surface wave.

It should also be pointed out that if the exact boundary conditions at the irregular surface are replaced by the approximate Kirchhoff type boundary conditions (Beckmann and Spizzichino 1963), the surface wave and lateral wave contributions are not taken into account.

To obtain the full wave solutions for transient excitations, analytical and numerical Fourier techniques are employed. For arbitrary time-varying excitations, the instantaneous expressions for the magnetic force  $\vec{J}_m$  for  $t \geq 0$  is

$$\vec{J}_m(\vec{r}, t) = \text{Re}[K f_s(t) \delta(x - x_0) \delta(y - y_0) \hat{a}_z] \quad (54)$$

where for convenience the excitation function  $f_s(t)$  is taken to be complex. The Fourier transform of  $f_s(t)$  is

$$F(\omega) = F[f_s(t)] = \int_0^\infty f_s(t) \exp(-i\omega t) dt \quad (55)$$

The transfer function  $H_s(\omega)$  for the scattered fields satisfies the relationship

$$H_s(\omega) = H_s^*(-\omega) \quad (56)$$

where  $*$  is the symbol for the complex conjugate. However, in general,  $F(\omega) \neq F^*(-\omega)$  since  $f_s(t)$  is complex. The transient response can be shown to be given by

$$\begin{aligned} h(\vec{r}, t) &= \text{Re} F^{-1} H_s(\omega) F(\omega) \\ &= \text{Re} \frac{1}{2\pi} \int_{-\infty}^{\infty} H_s(\omega) F(\omega) \exp(i\omega t) d\omega \\ &\equiv \text{Re} h_s(\vec{r}, t) \end{aligned} \quad (57)$$

In this work, LORAN C plus excitations are considered in detail. This pulse can be represented as a sum of three damped sinusoids (Johler and Horowitz 1973).

$$f_s(t) = \sum_{j=1}^3 A_j \exp(-\Gamma_j t) u(t) \quad (58)$$

in which  $u(t)$  is the unit step function.

$$\begin{aligned} A_1 &= i/2, \quad \Gamma_1 = c + i\omega_1, \quad \omega_1 = \omega_r \\ A_2 &= -i/4, \quad \Gamma_2 = c + i\omega_2, \quad \omega_2 = \omega_r + 2\omega_p \\ A_3 &= -i/4, \quad \Gamma_3 = c + i\omega_3, \quad \omega_3 = \omega_r - 2\omega_p \end{aligned} \quad (59)$$

Thus,

$$F(\omega) = \sum A_j / (s + \Gamma_j) \quad (60)$$

$$f_s(t) = i \exp(-ct) \sin^2 \omega_p t \exp(-i\omega_r t) \quad (61)$$

$$f(t) = |f_s(t)| \sin \omega_r t$$

and the envelope is

$$|f_s(t)| = \exp(-ct) \sin^2 \omega_p t \quad (62)$$

In general for dissipative media, the transfer functions  $H_s(\omega)$  are non-analytic. In these cases, the transient response  $h_s(\vec{r}, t)$  is expressed as follows (Bahar 1978a):

$$h_s(t) = h_L(t) + h_H(t) \quad (63)$$

where  $h_L(t)$  is the contribution to  $h_s(t)$  from the poles of  $F(s) = F(i\omega)$  (at  $s = -\Gamma_j$ ) and  $h_H(t)$  is the contribution to  $h_s(t)$  from the singularities of the transfer function  $H_s^*(s)$ . If  $H_s^*(s)$  does not possess singularities in the vicinity of the poles of  $F(s)$  it follows that  $|h_H(t)| \ll |h_L(t)|$ .

The term  $h_H(t)$  is significant only at the leading edge of the response. However, any distortion of the leading edge of the pulse due to medium effects could result in an effective delay in arrival time of the radio signal.

The instantaneous phase  $\phi(t)$  of the transient response is given by the phase of the complex response  $h_s(t)$ . Thus

$$\phi(t) = \arctan[\text{Im}\{f_s(t)\} / \text{Re}\{f_s(t)\}] \quad (64)$$

For the undistorted LORAN C pulse  $\phi(t) = -\omega_r t$ . However, in general, the phase deviation (or phase anomaly) for the transient response

$$\Delta\phi(t) = \phi(t) + \omega_r t \quad (65)$$

does not vanish. Rapid variations in  $\Delta\phi(t)$  occur at the leading edge of the transient response. Thus, examination of the terms  $h_H(t)$  and  $\Delta\phi(t)$  is also useful to determine medium effects on pulse delays.

For the illustrative examples presented in this section, the carrier frequency  $f_c = 10^5$  Hz,  $f_p = f_c/40$ , and  $c = 10f_p$ . The irregular boundary between the two media is given by

$$h(z) = h_{\max}[1 + \cos(\pi z/L)]/2, \quad -L \leq z \leq L \quad (66)$$

where  $2L = 10\lambda_{0c}$ ,  $k_0 h_{\max} = 2$ . The free-space wavelength at  $f_c$  is  $\lambda_{0c} = 3 \times 10^3$  m.

In Fig. 4, the instantaneous, specularly scattered radiation field is plotted as a function of  $t' = t - t_A$ , where  $t_A$  is the arrival time of the earliest response. Medium 1 ( $y < h$ ) is assumed to be highly conducting  $\epsilon_r = 100 - i10^6(\omega_c/\omega)$  and  $\theta_0' = \theta_0'' = 80^\circ$ . The envelopes of the scattered field and the undistorted LORAN C pulse are also given in Fig. 4. For convenience, both have been normalized to unity at their respective peaks.

For the case studied in Fig. 4, the pulse scattered in the specular direction undergoes very little distortion.

In Fig. 5, the non-specularly scattered radiation field is considered. For this case  $\epsilon_r = 10(1 - i\omega_c/\omega)$ ,  $\theta'' = 80^\circ$ , and  $\theta_0' = 50.8^\circ$ . This value of  $\theta_0'$  corresponds to the direction where the scattered field is at a minimum. In Fig. 5a in which the scattered instantaneous response and envelope are plotted, we note that the pulse is distorted. Furthermore, in Fig. 5b, the undulations in phase anomaly result from the fact that near the direction of a minima ( $\theta_0' = 50.8^\circ$ ) the scattered fields undergo destructive interference.

#### 4. INHOMOGENEOUS ANISOTROPIC IONOSPHERE EFFECTS ON NAVIGATIONAL AND POSITIONING SIGNALS

Numerical and analytical expressions are obtained for the reflected and transmitted horizontally and vertically polarized wave incident upon an inhomogeneous anisotropic layer such as the ionosphere. To this end, it is necessary to determine the transfer functions (the reflection and transmission coefficients) as a function of frequency and to apply Fourier transform techniques.

Full-wave solutions for the (steady-state) transfer functions are used. These full-wave solutions are based on a method which utilizes generalized characteristics vectors to determine the elements of a non-singular transformation matrix (Bahar 1976). This transformation matrix is used to convert Maxwell's equations for the transverse components of the electric and magnetic fields into a set of loosely coupled first-order differential equations. However these WKB solutions fail in critical coupling regions where the characteristic roots tend to merge. For these regions, numerous analytical and numerical techniques employing several special functions have been used to obtain suitable solutions for the fields. Special mathematical functions are not used here for the critical coupling regions. Instead, a suitable set of linearly independent wave amplitudes are derived through the use of the generalized characteristic vectors (Bahar 1976).

An approximate analytical expression for the transient response is also obtained through a suitable expansion of the transfer function about the carrier frequency. This analytical solution does not account for the singularities of the transfer function. However, since the third zero crossing of the response is regarded as the effective arrival time of the received signal, the analytical solutions can be used to determine the signal delay as well as the time of the third zero crossing. The fast Fourier transform (FFT) techniques are also used to determine the transient response, the amplitude distortion (due to the singularities of the transfer function), as well as the phase anomaly which influences the effective arrival time of the response.

Maxwell's equations for the transverse electromagnetic field components in horizontally stratified anisotropic media, such as the ionosphere can be expressed in separable form as follows (Budden 1962):

$$e' = \frac{de}{du} = -\frac{1}{ik} \frac{de}{dz} = T'e, \quad (67)$$

in which

$$e = \begin{bmatrix} E_x \\ -E_y \\ H_x \\ H_y \end{bmatrix}, \quad (68)$$

$$u - u_0 = -ik(z - z_0), \quad (69)$$

and  $k_0 = \omega(\mu_0\epsilon_0)^{1/2}$  is the free-space wavenumber. The elements  $T_{ij}$  of the  $4 \times 4$  dimensionless matrix  $T$  which are related to the susceptibility matrix  $M$  are in general also function of  $u$  (Budden 1962). The  $z$  axis is normal to the horizontally stratified media, and the direction of the incident wave normal in free space is

$$u = (sa_r + ca_r), \quad (70)$$

where  $s$  and  $c$  are the sine and cosine of the angle of incidence  $\theta$  for propagating waves. Thus, the fields are independent of  $y$ , and, assuming time harmonic excitations, the factor  $\exp(-iskx)\exp(i\omega t)$  is suppressed throughout this work. The transverse electric field components are  $E_x$  and  $E_y$ , while  $H_x$  and  $H_y$  are the transverse components of the normalized magnetic field  $\eta H$  where  $\eta = (\mu_0/\epsilon_0)^{1/2}$  is the free-space wave impedance, and  $H$  is the magnetic field.

When the parameters of the susceptibility matrix  $M$  (plasma frequency  $f_p$ , collision frequency  $\nu$ , gyrofrequency  $f_H$ , and the direction of the earth's magnetic field  $\hat{a}_H$ ) are slowly varying functions of  $z$  and the medium is devoid of critical coupling regions, the matrix transformation

$$e = Sf \quad (71)$$

can be used to convert Maxwell's equations (67) into a set of loosely coupled equations for the upward and downward propagating ordinary and extraordinary wave amplitudes  $f_i$  ( $i = 1, 2, 3, 4$ ):

$$f' = \Delta f + \Gamma f \quad (72)$$

in which

$$\Delta = S^{-1}TS, \quad \Gamma = S^{-1}S' \quad (73)$$

The matrix transformation  $S^{-1}TS$  yields a diagonal matrix  $\Delta$  whose elements are the characteristic values of the matrix  $T$ . These characteristic values  $q_i$  are solutions to the Booker quartic (Budden 1962):

$$\det(T - q_i I) \equiv |T - q_i I| = 0 \quad (74)$$

in which  $I$  is the identity matrix of rank  $n = 4$ . The columns of the  $4 \times 4$  matrix  $S$  are the four characteristic vectors that satisfy

$$TS' = q_i S' \quad (75)$$

Since the medium is varying  $T$ ,  $q_i$ , and  $S'$  are functions of  $z$  in (72). To solve (72) numerically for the region  $-\Delta z/2 \leq z - z_c \leq \Delta z/2$ , it is convenient to evaluate  $S'$  using (75) at  $z = z_c$  for the entire partition of width  $\Delta z$ . Thus for each partition  $|z - z_c| \leq \Delta z/2$ , (72) is replaced by

$$\begin{aligned} f' &= S^{-1}TSf \equiv S^{-1}(T_c + T_\delta)Sf \\ &\equiv (\Delta_c + \gamma)f \equiv Cf \end{aligned} \quad (76)$$

in which

$$T_c = T(z_c), \quad \Delta_c = \Delta(z_c), \quad \text{and} \quad \gamma = S^{-1}T_0S \quad (77)$$

The elements  $\gamma_{mn}$  of the coupling matrix  $\gamma$  vanish at  $z = z_c$  since  $T_\delta = T - T_c$  vanishes at the center of the partition of width  $\Delta z$ . When the fourth-order Runge-Kutta method (Abramowitz and Stegun 1964) is used to solve (76), the partition size must be chosen such that  $|C_{mn}k\Delta z| \ll 1$ , where  $C_{mn}$  are the elements of the matrix  $C$ .

For critical coupling regions where the characteristic values  $q_i$  tend to merge, the transformation matrix  $S$  obtained by using (75) becomes near singular, and the coupling coefficients  $\gamma_{mn}$  become very large. In these regions, the interaction between the characteristic waves  $f_i$  constitute large reflections or strong coupling between the ordinary and extraordinary waves traveling in the same direction. Thus, the transformation based on (75) is not suitable for critical coupling regions. For these regions, it has been shown that it is possible to construct a non-singular transformation matrix  $S$  by using generalized characteristic vectors (Bahar 1976). Thus for instance, if for the matrix  $T = T_0$  the characteristic equation  $|T_0 - qI| = 0$  has a root  $q_i$  of multiplicity  $m_i$ , it has been shown that, associated with this root, there exists  $m_i$  linearly independent generalized characteristic vectors. In this case, even though  $T_0$  and the generalized characteristic vectors are independent of  $z$ , the transformation matrix  $S$  is  $z$ -dependent. It can be shown that the non-singular matrix  $S$  satisfies the equation (Bahar 1976)

$$S^{-1}(T_0S - S') = \Delta \quad (78)$$

in which  $\Delta$  is also a diagonal matrix. Thus for critical coupling regions, where the characteristic values  $q_i$  tend to merge, (67) can be transformed to the following matrix equation:

$$\begin{aligned} f' &= S^{-1}(TS - S')f = S^{-1}[(T_0 + T_\delta)S - S']f \\ &= (\Delta + \gamma)f \equiv Cf \end{aligned} \quad (79)$$

in which

$$T_\delta = T - T_0 \quad \text{and} \quad \gamma = S^{-1}T_\delta S \quad (80)$$

Hence, at a critical coupling point where  $T = T_0$ , the elements of the coupling matrix  $\gamma_{mn}$  vanish.

Since the transverse components of the electromagnetic fields are continuous for sourceless regions, at the interface ( $z = z_T$ ) between two adjacent partitions the continuity condition is

$$e(z_T) = S^-f(z_T^-) = S^+f(z_T^+), \quad (81)$$

where  $S^-$  and  $S^+$  are the transformation matrices for  $z \approx z_T^-$  and  $z \approx z_T^+$ , respectively, and  $z_T^\pm = \lim_{\delta \rightarrow 0} z_T \pm \delta$  ( $\delta > 0$ ).

The transfer functions for an inhomogeneous anisotropic dielectric layer are the horizontally and vertically polarized reflected and transmitted wave amplitudes excited by a horizontally or a vertically polarized incident wave of unit amplitude. Thus for instance, the reflection and transmission coefficients  $R_{VH}(\omega)$  and  $T_{VH}(\omega)$  are the vertically polarized reflected and transmitted wave amplitudes, respectively, due to an incident horizontally polarized wave of unit amplitude. For a dielectric layer of thickness  $L$  ( $0 < z < L$ ), there are eight transfer function  $R_{PQ}(\omega)$  and  $T_{PQ}(\omega)$  ( $P, Q = V$  or  $H$ ) for waves incident from below the layer ( $z < 0$ ) and eight transfer functions for waves incident from above ( $z > L$ ). They are related through the adjoint reciprocity relationship in electromagnetic theory (D. M. Kerns 1976). To obtain

these transfer functions, it is necessary to numerically solve the first-order coupled differential equations (76) or (79) in conjunction with the continuity conditions (81) at the interface between two adjacent partitions of width  $\Delta z$ . For free space above and below the anisotropic dielectric layer, the transformation matrix  $S$  can be written as (Bahar 1976)

$$S = \begin{bmatrix} 0 & 0 & c & -c \\ 1 & 1 & 0 & 0 \\ c & -c & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \quad (82)$$

Thus for an incident horizontally polarized wave of amplitude  $f_0^{IH}$  at  $z = 0$ , the value of the wave amplitudes at  $z = 0^-$  and  $z = L^+$  are  $f^{IH}$  and  $f^{TH}$ , respectively, where

$$f^{IH} = f_0^{IH} \begin{bmatrix} 1 \\ R_{HH} \\ 0 \\ R_{VH} \end{bmatrix} \quad (83)$$

$$f^{TH} = f_0^{IH} \begin{bmatrix} T_{HH} \\ 0 \\ T_{VH} \\ 0 \end{bmatrix} \quad (84)$$

Similarly, for an incident vertically polarized wave of amplitude  $f_0^{IV}$  at  $z = 0$ , the values for the wave amplitude at  $z = 0^-$  and  $z = L^+$  are  $f^{IV}$  and  $f^{TV}$ , respectively, where

$$f^{IV} = f_0^{IV} \begin{bmatrix} 0 \\ R_{HV} \\ 1 \\ R_{VV} \end{bmatrix}, \quad (85)$$

$$f^{TV} = f_0^{IV} \begin{bmatrix} T_{HV} \\ 0 \\ T_{VV} \\ 0 \end{bmatrix} \quad (86)$$

Since the values of the wave amplitudes  $f$  are unknown initially at either boundary  $z = 0^-$  or  $z = L^+$ , we first solve (79) subject to the two separate boundary conditions at  $z = L^+$ :

$$g^T = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad h^T = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad (87)$$

Denoting these solutions as  $g(z)$  and  $h(z)$ , (whose elements are  $g_i$  and  $h_i$ ,  $i = 1, \dots, 4$ ) respectively, and setting  $g(0^-) = g^I$ ,  $h(0^-) = h^I$ , the wave amplitudes  $f(z)$  can be expressed as a superposition of these two independent solutions. Thus

$$f^H(z) = g(z) + \alpha^H h(z) \quad (88)$$

$$f^V(z) = \alpha^V g(z) + h(z) \quad (89)$$

Equating (88) for  $z = 0^-$  and  $z = L^-$  to (83) and (84), respectively, we get

$$\alpha^H = -g_3^I/h_3^I, \quad f_0^{IH} = g_1^I + \alpha^H h_1^I \quad (90)$$

$$\begin{aligned} \begin{bmatrix} 1 \\ R_{HH} \\ 0 \\ R_{VH} \end{bmatrix} &= \frac{1}{f_0^{IH}} (g^I + \alpha^H h^I), \\ \begin{bmatrix} T_{HH} \\ 0 \\ T_{VH} \\ 0 \end{bmatrix} &= \frac{1}{f_0^{IH}} \begin{bmatrix} 1 \\ 0 \\ \alpha^H \\ 0 \end{bmatrix} \end{aligned} \quad (91)$$

Similarly, equating (89) for  $z = 0^-$  and  $z = L^-$  to (85) and (86), respectively, we get

$$\alpha^V = -h_1^I/g_1^I, \quad f_0^{IV} = \alpha^V g_3^I + h_3^I \quad (92)$$

$$\begin{bmatrix} 0 \\ R_{HV} \\ 1 \\ R_{VV} \end{bmatrix} = \frac{1}{f_0^{IV}} (\alpha^V g^I + h^I), \quad \begin{bmatrix} T_{HV} \\ 0 \\ T_{VV} \\ 0 \end{bmatrix} = \frac{1}{f_0^{IV}} \begin{bmatrix} \alpha^V \\ 0 \\ 1 \\ 1 \end{bmatrix} \quad (93)$$

For low frequencies where the width  $L$  of the anisotropic dielectric considered is of the order of several wavelengths, the numerical solutions to (79) can be obtained using a fourth-order Runge-Kutta method (Abramowitz and Stegun 1964). Typically, the number of partitions needed is  $20L/\lambda > N > 10L/\lambda$ . For high frequencies  $f \gg f_p$  and  $L/\lambda \gg 1$ , the Runge-Kutta method is not suitable, and it is more convenient to convert the coupled differential equations (79) into a set of coupled integral equations (Bahar and Agrawal 1979).

Numerical as well as analytical techniques are used to determine the transient response to navigational signals. The LORAN C excitation  $f(t)$  (considered here in detail) is expressed in terms of the complex functions  $f_p(t)$  as in Section 3.

The transfer functions  $R_{PQ}(\omega)$  and  $T_{PQ}(\omega)$  satisfy the relationship

$$R_{PQ}(\omega) = R_{PQ}^*(-\omega) \text{ and } T_{PQ}(\omega) = T_{PQ}^*(-\omega) \quad (94)$$

where the asterisk denotes the complex conjugate. Thus it can be shown that the full-wave expression for the transient response is the real part of the inverse Fourier transform:

$$\begin{aligned} \text{Re } e_s(t) &= \text{Re } F^{-1}[F(\omega)R(\omega)] \\ &= \text{Re } \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)R(\omega) \exp(i\omega t) d\omega \\ &= \text{Re}[e_R(t) + ie_I(t)] = e_R(t) \end{aligned} \quad (95)$$

in which  $R(\omega)$  is any one of the eight transfer functions  $R_{PQ}(\omega)$  and  $T_{PQ}(\omega)$ . To obtain an approximate analytical solution for  $e(t)$ , we express the transform  $R(\omega)$  as follows:

$$R(\omega) = |R(\omega)| \exp[i\phi(\omega)] \cong R(-\omega_c) \exp[-i(\omega + \omega_c)\tau_R], \quad (96)$$

in which

$$R(-\omega_c) = |R(-\omega_c)| e^{i\phi(-\omega_c)} = R^*(\omega_c) \quad (97)$$

is the value of  $R_{PQ}$  at  $\omega = -\omega_c$ , and the time delay  $\tau_R$  is

$$\tau_R = -\frac{d}{d\omega} \phi(\omega) |_{\omega=\pm\omega_c} \quad (98)$$

Thus on substituting (95) into (61), we get

$$\begin{aligned} e_s(t) &\cong i \exp[-c_1(t - \tau_R)] \sin^2[\omega_p(t - \tau_R)] \\ &\quad \cdot \exp[-i\{\omega_c t - \phi(-\omega_c)\}] U(t - \tau_R) \end{aligned} \quad (99)$$

Noting that  $\phi(-\omega_c) = -\phi(\omega_c)$  and setting  $t' = t - \tau_R$ , (99) can also be written as

$$\begin{aligned} e_s(t') &= i \exp[-c_1 t'] \sin^2 \omega_p t' \\ &\quad \cdot \exp[-i\{\omega_c(t' + \tau_R) + \phi(\omega_c)\}] U(t') \end{aligned} \quad (100)$$

Thus in the above analytical expression, the envelope of the response with respect to the shift time  $t'$  is the same as the

envelope of the incident LORAN C pulse; however, the instantaneous phase is  $\omega_c(t' + \tau_R) + \phi(\omega_c)$ . The approximate analytical expression (100) for (95) only accounts for the residues at the poles of the transform  $F(\omega)$ . However, the expression (95) accounts for singularities of both  $F(\omega)$  and  $R(\omega)$ . The singularities of  $R(\omega)$  account for the distortion of the transient response which is expressed as follows:

$$\begin{aligned} \Delta e(t') &= \text{Re } \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) [R(\omega) \exp(i\omega \tau_R) \\ &\quad - R^*(\omega_c) \exp(-i\omega_c \tau_R)] \exp(i\omega t) d\omega \\ &\equiv \Delta e_R(t') + i \Delta e_I(t') \end{aligned} \quad (101)$$

Since an analytical expression for  $R(\omega)$  is not known for arbitrary anisotropic media, both (95) and (101) can only be evaluated numerically. The FFT techniques are found suitable to evaluate these integrals. The instantaneous phase of the transient response (95) is given by

$$\phi(t') = \tan^{-1}[e_I(t')/e_R(t')] \quad (102)$$

Thus the phase anomaly of the instantaneous response is

$$\Delta \phi(t') = \phi(t') + \omega_c(t' + \tau_R) + \phi(\omega_c) \quad (103)$$

For the illustrative examples presented in this section, the carrier frequency  $f_c = 10^5$  Hz,  $f_p = f_c/10$ , and  $c_1 = 10f_p$ . The inhomogeneous anisotropic ionosphere is characterized by the plasma, collision, and gyrofrequencies  $\omega_p$ ,  $\nu$ , and  $\omega_H$ , respectively. The inhomogeneous plasma frequency  $f_p$  is assumed to be given by

$$\begin{aligned} X(z) &= \left( \frac{f_p(z)}{f} \right)^2 = \frac{X_1}{2} [1 - \cos(\pi z/L)], \\ 0 < z < L &= \lambda_c/2 \end{aligned} \quad (104)$$

where

$$X_1 = X(L) = (\omega_p(L)/\omega)^2 = (\lambda/\lambda_p(L))^2 \quad (105)$$

and

$$\begin{aligned} \lambda_p(L) &= 3.5 \times 10^3 \text{ m} = c_0/f_p(L) \\ \lambda_c &= c_0/f_c = \frac{(c_0 \mu_0)^{-1/2}}{f} \end{aligned} \quad (106)$$

The collision frequency is given by

$$Z(z) = \nu(z)/\omega = Z_1 \exp(-\gamma z), \quad 0 < z < L \quad (107)$$

where

$$Z_1 = Z(0) = \frac{\nu(0)}{\omega} = \lambda/\lambda_\nu(0) \quad (108)$$

and

$$\lambda_\nu(0) = 2\pi c_0/\nu = 3 \times 10^4 \text{ m}, \quad \gamma = 2 \times 10^{-4} \text{ m}^{-1} \quad (109)$$

The gyrofrequency  $f_H$  is given by

$$Y(z) = \frac{f_H}{f} = \lambda/\lambda_H \quad (110)$$

where

$$\lambda_H = c_0/f_H = 3 \times 10^2 \text{ m} \quad (111)$$

The direction of the earth's magnetic field  $\hat{a}_H$  is given by

$$Y = (0.5\hat{a}_r + \sqrt{0.5}\hat{a}_\theta + 0.5\hat{a}_z)Y = a_H Y \quad (112)$$

In Figs. 6 and 7, the transfer functions  $R_{HH}(\omega)$ , and  $T_{HH}(\omega)$  are given for waves incident from below the anisotropic layer at an angle  $\theta = 45^\circ$ ; thus the direction of the wave normal is  $\hat{n} = (\hat{a}_x + \hat{a}_y)/\sqrt{2}$ . The value of the transfer functions at  $\omega = \omega_c/2$ ,  $\omega = \omega_c$ , and  $\omega = 2\omega_c$  are indicated by symbols  $\square$ ,  $\Delta$ , and  $*$ , respectively.

In Figs 8 and 9, the normalized values of  $e_{RN}(t')$  and its envelope  $|e_{RN}(t')|$  are given for the transfer functions  $R_{HH}(\omega)$  and  $R_{VH}(\omega)$ . These functions are evaluated using the FFT algorithm as well as the analytic expression (99). The difference between the numerical FFT results and the analytical results  $\Delta e(t')$  (101) is negligible for  $t > 3/f_0$ . The effect of the medium is to cause the delay  $\tau_R$  and the shift in the third zero crossing of approximately half the period  $1/f_c$ . The delay  $\tau_R$  is not shown in the figures since  $e_{RN}$  is plotted as a function of the shifted time scale  $t' = t - \tau_R$ . For the cases considered in Figs. 8 and 9,  $\tau_R$  is  $6.666 \times 10^{-6}$ s and  $6.660 \times 10^{-6}$ s, respectively. There is excellent agreement between the value for  $\tau_R$  obtained from (98) and the numerical results obtained by using the FFT algorithm in Figs. 8 and 9.

**ACKNOWLEDGEMENTS:** The author wishes to thank the organizing committee of the AGARD-EPP Symposium on Navigation and Positioning Systems for support to attend the meeting and is grateful to Dr. J. H. Richter for his encouragement. The manuscript was prepared by Dr. M. El-Shenawee and L. Villalobos.

## 5. REFERENCES

- Abramowitz, M. and I. A. Stegun, 1964. *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, Applied Math Series, 55, Nat. Bur. of Stand. Washington, D.C.
- Bahar E., 1972, Radio wave propagation over a rough variable impedance boundary, 1, Full wave analysis. *IEEE Trans. Antennas Propagat.*, AP-20 (3), 362-368.
- Bahar E., 1974, Depolarization in nonuniform multilayered structures - full wave solutions, *Journal of Mathematical Physics*, 15, 202-208.
- Bahar, E., 1976, Generalized characteristic functions for simultaneous linear differential equations with variable coefficients applied to propagation in inhomogeneous anisotropic media. *Canadian Journal of Physics*, 54, 301-306.
- Bahar, E., 1977, Coupling between guided surface waves, lateral waves, and the radiation fields by rough surfaces - Full wave solutions. *IEEE Trans. Microwave Theory Tech.* MTT-25 (11), 923-931.
- Bahar, E., 1978a, Transient electromagnetic response from irregular models of the earth's surface. *Radio Science*, 13, 345-355.
- Bahar, E., 1978b, Full wave and physical optics solutions for scattered radiation fields for rough surfaces - Energy and reciprocity relationships. *IEEE Trans. Antennas Propagat.*, AP-26 (4), 603-614.
- Bahar, E. and B. S. Agrawal, 1979a, Generalized characteristic functions applied to propagation in bounded inhomogeneous anisotropic media. *Journal of Atmospheric and Terrestrial Physics*, 41, 565-578.
- Bahar, E. and B. S. Agrawal, 1979b, Ionosphere effects on navigational and positioning signals - full wave solutions. *IEEE Trans. on Antennas and Propagation*, AP-27, 225-232.
- Bahar, E. and B. S. Agrawal, 1979c, Distortion and depolarization by the ionosphere of L band signals coded by phase reversals: full wave solutions. *Radio Science*, 14, 843-853.
- Bahar, E. 1980, Computations of the transmission and reflection scattering coefficients in an irregular spheroidal model of the earth-ionosphere waveguide. *Radio Science*, 15, 987-1000.
- Bahar, E. and M. A. Fitzwater, 1981, Numerical technique to trace the loci of the complex roots of characteristic equations in mathematical physics. *SIAM Journal on Scientific and Statistical Computation*, 2, 389-403
- Beckmann, P. and A. Spizzichino, 1963. *The Scattering of Electromagnetic Waves from Rough Surfaces*, Macmillan, New York.
- Brigham, E. O., 1974. *The Fast Fourier Transform*, Prentice-Hall, Englewood Cliffs, NJ.
- Budden, K. G., 1962, *Radio Waves in the Ionosphere*, Cambridge University Press.
- Johler, J. R. and S. Horowitz, 1973, Propagation of LORAN ground and ionospheric wave pulses. *Office of Telecommunications Rep. 73-20*, US Government Printing Office, Washington, DC 20402.
- Kerns, D. M., 1976, Plane-wave scattering matrix theory of antennas and antenna-antenna interactions: Formulation and applications. *National Technical Information Service*, Springfield, VA, NBSIR-75-824.
- Wait, J. R., 1962. *Electromagnetic Waves in Stratified Media*, Pergamon, New York.
- Watson, G. N., 1918, The diffraction of electric waves by the earth. *Proc. Roy. Soc. Ser. A*, 95, 83-99.



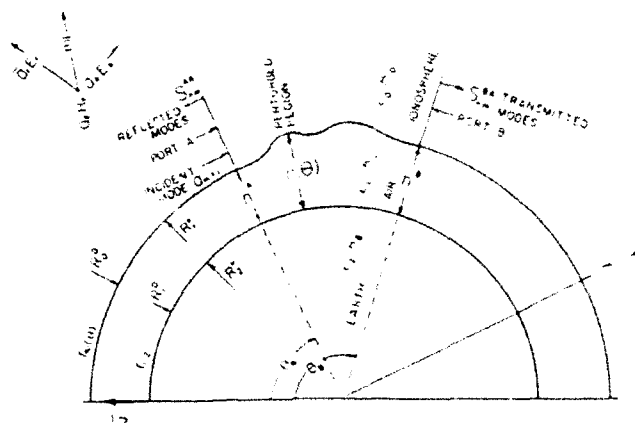


Fig. 1. Irregular multilayered spheroidal model of the earth-ionosphere waveguide.

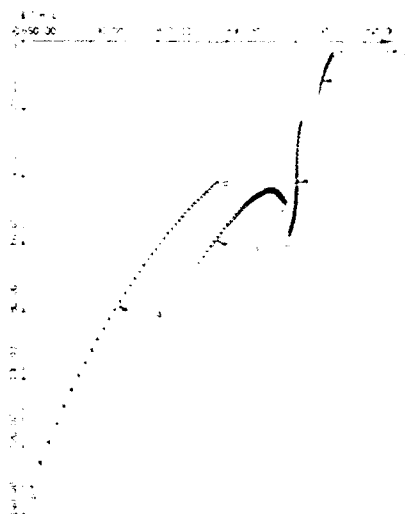


Fig. 2. Loci of the roots  $\nu_n$  for  $n = 1, 2, 3, 4$  as  $h$  varies from 60 km to 90 km.

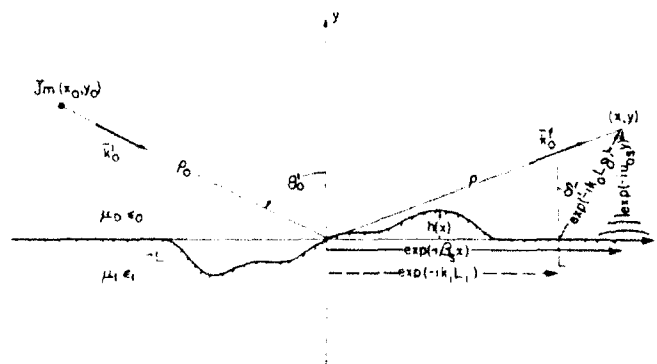


Fig. 3. The scattered radiation fields, surface waves and lateral waves due to incident plane waves over rough surfaces.

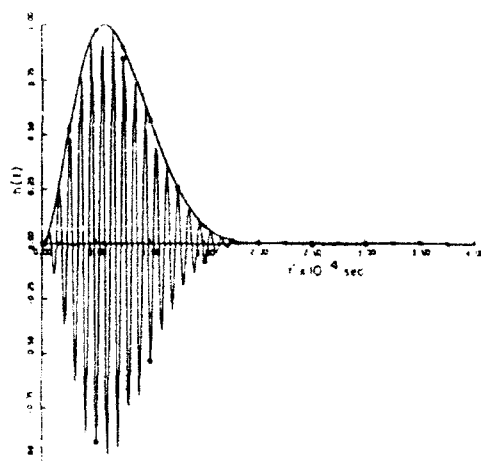


Fig. 4. The scattered radiation field for  $\epsilon_r = 100 - i 10^{-5} \frac{\omega}{\omega_0}$  and  $\theta_0^i = \theta_0^f = 80^\circ$ . The instantaneous response and the envelopes.

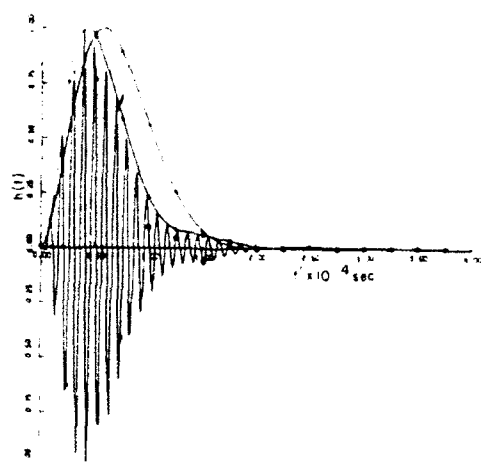


Fig. 5a. The scattered radiation field for  $\epsilon_r = 10(1 - i \frac{\omega}{\omega_0})$ ,  $\theta_0^i = 80^\circ$  and  $\theta_0^f = 50.8^\circ$ . The instantaneous response and the envelopes.

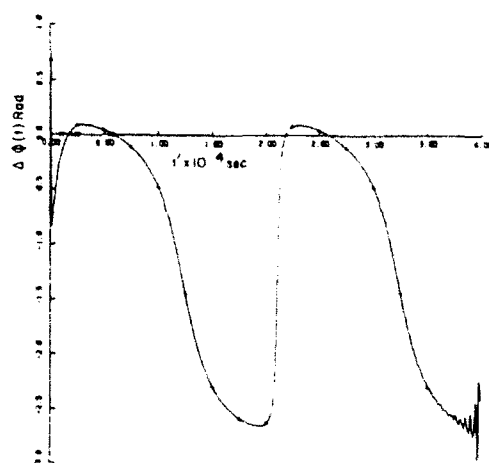


Fig. 5b. The scattered radiation field for  $\epsilon_r = 10(1 - i \frac{\omega}{\omega_0})$ ,  $\theta_0^i = 80^\circ$  and  $\theta_0^f = 50.8^\circ$ . The phase anomaly  $\Delta(\phi)$ .



Fig. 6. Reflection coefficient for horizontally polarized waves  $R_{HH}(\omega)$ .

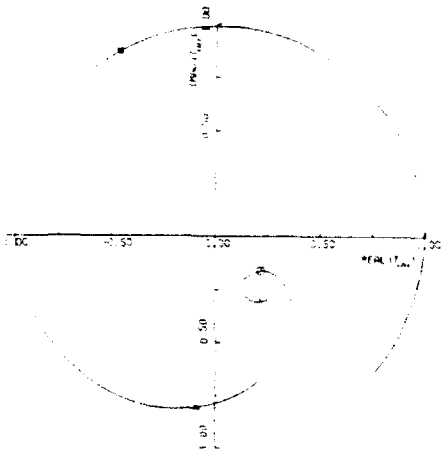


Fig. 7. Transmission coefficient for horizontally polarized waves  $T_{HH}(\omega)$ .

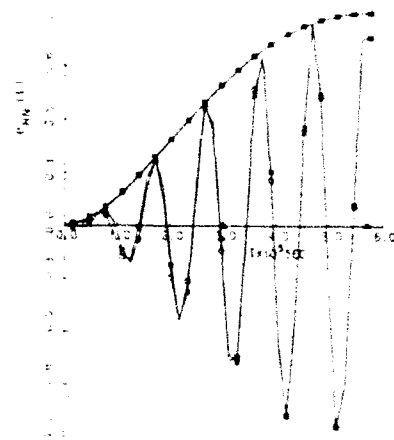


Fig. 8. Normalized transient response  $e_{RN}(t')$  and its envelope  $|e_{RN}(t')|$  corresponding to the transfer function  $R_{HH}(\omega)$ . (\*) analytic, (□) FFT.

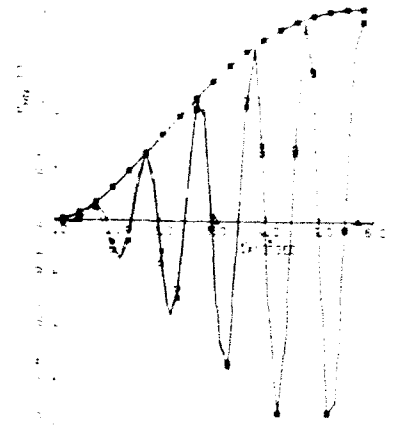


Fig. 9. Normalized transient response  $e_{RV}(t')$  and its envelope  $|e_{RV}(t')|$  corresponding to the transfer function  $R_{VH}(\omega)$ . (\*) analytic, (□) FFT.

TABLE I. COMPOSITE SCATTERING MATRIX

$$\begin{bmatrix} B^+ \\ B^- \end{bmatrix} = \begin{bmatrix} S^{++} & S^{+-} \\ S^{-+} & S^{--} \end{bmatrix} \cdot \begin{bmatrix} A^+ \\ A^- \end{bmatrix}$$

0.174D - 03 0.178	0.214D - 03 2.536	0.570D - 03 0.455	0.146D - 02 2.638	0.737D - 00 1.461	0.545D - 00 1.274	0.822D - 01 0.260	0.416D - 01 2.752
0.214D - 03 2.836	0.276D - 03 0.616	0.735D - 03 2.186	0.184D - 02 0.597	0.195D - 00 1.323	0.288D - 00 2.937	0.690D - 00 1.028	0.855D - 01 0.350
0.570D - 03 0.455	0.735D - 03 2.486	0.269D - 02 0.792	0.457D - 02 2.396	0.231D - 00 1.410	0.239D - 00 1.781	0.134D - 01 2.600	0.648D - 00 0.434
0.146D - 02 2.638	0.184D - 02 0.597	0.457D - 02 2.396	0.127D - 01 0.795	0.408D - 00 2.575	0.458D - 00 0.191	0.192D - 01 2.471	0.261D - 01 0.486
0.737D - 00 1.461	0.195D - 00 1.323	0.231D - 00 1.410	0.408D - 00 2.575	0.462D - 02 0.566	0.375D - 02 2.131	0.693D - 01 0.695	0.355D - 02 0.063
0.545D - 00 1.274	0.288D - 00 2.937	0.239D - 00 1.781	0.458D - 00 0.191	0.375D - 02 2.131	0.304D - 02 1.463	0.629D - 01 1.868	0.129D - 02 2.767
0.822D - 01 0.260	0.690D - 00 1.028	0.134D - 01 2.600	0.192D - 01 2.471	0.693D - 01 0.695	0.629D - 01 1.868	0.292D - 01 1.683	0.477D - 01 0.539
0.416D - 01 2.752	0.855D - 01 0.350	0.648D - 00 0.434	0.261D - 01 0.486	0.355D - 02 0.063	0.129D - 02 2.767	0.477D - 01 0.539	0.419D - 01 1.486

The transmission and reflection scattering coefficients magnitude is the top number of each pair, and the phase, in radians, is the bottom number.

## INFLUENCE DE LA PROPAGATION SUR LA GONIOMETRIE HAUTE FREQUENCE ET LA LOCALISATION A STATION UNIQUE

V. BALTAZART, L. BERTEL  
UNIVERSITE DE RENNES I - LABORATOIRE RADIOCOMMUNICATIONS, URA 834  
CAMPUS DE BEAULIEU - 35042 RENNES CEDEX - FRANCE

R. FLEURY  
CNET - LAB/PT/GER  
B.P. 40 - 22301 LANNION CEDEX - FRANCE

### RESUME

Nous examinons dans cet article les effets de propagation sur les systèmes de Localisation à Station Unique de la gamme décimétrique.

Nous considérons tout d'abord, le problème direct qui consiste à estimer les angles d'arrivée des différents plans d'ondes incidents. Pour une situation du milieu ionosphérique donnée, nous déterminons pour cela les angles d'incidence de l'ensemble des trajets et modes de propagation. Le modèle simplifié que nous développons permet ensuite la simulation d'un comportement d'un radiogoniomètre.

On donne un aperçu des performances des systèmes basés sur des techniques d'interférométrie, de traitement de données et d'analyse spatiotemporelle. La proximité des modes de propagation nécessite dans certains cas d'ajouter une procédure de filtrage de polarisation.

Le problème inverse consiste à localiser un émetteur à l'aide des résultats de goniométrie précédents. Nous voyons l'effet d'une incertitude sur le profil, avant d'introduire deux techniques de LSU : l'une traite le problème à partir des résultats de sondage vertical au point de réception, la seconde réactualise en oblique les paramètres essentiels du profil. Un exemple d'influence de l'inclinaison des couches est également donné.

### ABSTRACT

This paper deals with some propagation effects which can influence Single Station Location systems.

We consider first, the direct problem which consists of estimating several directions of arrival of the incoming waves. For this purpose and for a given state of the ionospheric medium, the angles of all propagation paths and modes are determined. In a second step, the simplified model of waves coming from the ionosphere which has been developed before, permits us to simulate the behaviour of radiogoniometers.

We analyse the performance of goniometric systems based upon interferometry, data processing and space-time processing. Proximity in such propagation modes leads to the addition of a polarisation filtering procedure.

The problem in reverse consists of locating the emitter thanks to the previous radiogoniometry results. We look at the effect of an imperfect knowledge of the profile, before introducing two SSL methods : the first dealing with vertical sounding results ; the second works from a partial updating profile. We give also an example of an influence of tilt.

### MOTS-CLES

Propagation ionosphérique - multitrajets - anisotropie - doppler - trace de rayon - modèle MQP - radiogoniométrie - interférométrie - filtrage de polarisation - localisation à station unique - tilt - réactualisation.

### INTRODUCTION

Les techniques de radiolocalisation à station unique ont toujours suscité un grand intérêt lié à l'exploitation aisée des résultats espérés. Ces techniques s'appuient sur :

- une procédure d'obtention des angles d'arrivée, (élévation et azimut)
- une recherche de la source d'émission basée sur une modélisation à une ou plusieurs dimensions du profil de densité électronique associée à une technique de tracé de rayon réel ou équivalent.

A l'inverse de la plupart des publications qui sont citées ici par la suite, cet article aborde les deux aspects du problème en parallèle.

Dans une première partie consacrée au problème direct (détermination des angles d'arrivée), on suppose connu le comportement du milieu ionosphérique dans l'espace et le temps. Les paramètres utilisés sont déduits de prévisions de propagation.

Ainsi nous déterminons d'abord, pour une liaison donnée, le nombre de trajets, de modes de propagation et la polarisation des diverses ondes incidentes au système. Ceci conduit, à mettre en évidence les contraintes imposées aux radiogoniomètres HF, et aussi à définir des limites d'utilisation.

La modélisation des signaux reçus en vue de tester les divers types de radiogoniomètres, revêt une grande importance. Nous décrivons le modèle de signal utilisé incluant la polarisation des ondes incidentes, la réponse des antennes à ces ondes et les dopplers intermodes.

Le comportement de certains systèmes existants est analysé ; le modèle conduit également à associer à un goniomètre HF testé récemment, une technique de réception en diversité de polarisation.

Enfin, dans une dernière partie, quelques aspects du problème inverse sont abordés : influence des paramètres du profil de densité, de l'inclinaison des couches, nécessité de disposer d'informations sur le profil de densité électronique.

## 1. LIAISONS POINT A POINT, DETERMINATION DU NOMBRE DE TRAJETS ET DE MODES

Le premier problème qui se pose lorsqu'on veut simuler le comportement d'un système de radiogoniométrie HF est de savoir comment s'effectuent les liaisons point à point par voie ionosphérique. Ceci nécessite deux étapes :

- la définition d'un profil de densité électronique
- une technique de détermination des angles d'arrivée.

### 1.1 Profil de densité électronique

Nous avons choisi un profil de densité prenant en compte les différentes couches de densité électronique (E, F1, F2). Parmi les différents modèles disponibles nous avons choisi le modèle multiquasiparabolique de Hill [1] représentant le profil de densité  $N(r)$  à partir de morceaux de quasi-paraboles. Ce modèle (MQP) est actuellement utilisé par de nombreuses équipes [2, 3]. On a ainsi :

$$N(r) = \sum_{i=1}^P N_{mi} \left[ 1 \pm \left[ \frac{r - r_{mi}}{y_{mi}} \right]^2 \left[ \frac{r_{bi}}{r} \right]^2 \right]$$

avec :

- $r$  : distance point ionosphérique considéré - centre de la terre
- $N_{mi}$  : densité au maximum (ou minimum) de la couche parabolique  $i$
- $r_{mi}$  : valeur de  $r$  pour  $N(r) = N_{mi}$
- $r_{bi}$  : valeur de  $r$  à la base de la couche considérée
- $y_{mi}$  : demi-épaisseur pour la couche quasiparabolique considérée
- $P$  : nombre de morceaux de quasiparaboles

Le modèle est construit avec 3 couches (E, F1, F2) (figure 1) et une vallée entre E et F1. Nous plaçons entre deux couches ainsi définies des couches de liaison également de type quasiparaboliques. Une telle couche peut également être utilisée pour assurer une continuité entre le bas de l'ionosphère (pris à 80 km) et la couche E. Ceci conduit à 7 ou 8 morceaux de quasiparaboles, selon qu'on prenne en compte cette dernière couche ou non. Les paramètres de la vallée sont déduits de ceux de la couche E suivant la technique utilisée dans le logiciel d'inversion des ionogrammes verticaux POLAN [4]. Les profils ainsi définis rendent bien compte des observations.

Dans nos simulations, nous calculons les principaux paramètres par prévisions de propagation. Nous déterminons ainsi les fréquences critiques des couches E, F1, F2, l'altitude et la demi-épaisseur de la couche F2. Les autres paramètres sont estimés. A l'image de la vallée, nous nous donnons la possibilité de remplacer ces valeurs par celles déduites de sondages verticaux et par application du logiciel POLAN.

### 1.2 Techniques de détermination des angles d'arrivée lorsque la position de l'émetteur est connue

Suivant que l'on tienne compte ou non du champ magnétique terrestre, la technique mise en oeuvre pour déterminer les angles d'arrivée sera différente. Dans le cas où le milieu est considéré comme isotrope, une méthode simple est décrite dans [5]. Cette méthode s'appuie sur :

- l'avantage de la représentation du profil par morceaux de quasiparaboles conduisant à une expression analytique de la distance de saut en fonction des paramètres du profil obtenus à partir de prévisions de propagation ou de mesures, de la fréquence et de l'angle d'élévation. Elle conduit à la détermination des angles d'arrivée pour la liaison considérée.

ALTITUDE (km)

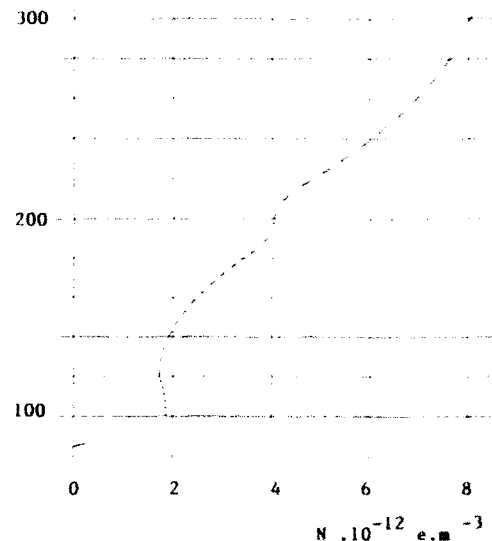


FIGURE 1 - PROFIL DE DENSITE ELECTRONIQUE  
ELECTRONIC DENSITY PROFILE

- l'application de la loi de Bouguer, l'angle d'élévation étant calculé à partir de :

$$E_k = \text{Arc cos} \left[ \frac{r_{tk} n(r_{tk})}{r_0} \right]$$

- $E_k$  : angle d'arrivée correspondant au mode  $k$
- $r_{tk}$  : distance centre de la terre-point de réflexion déterminé par itération
- $n(r_{tk})$  : indice de phase pour  $r = r_{tk}$

On déduit des valeurs de  $E_k$  et  $r_{tk}$  les temps de groupe et de phase.

Dans le cas où on tient compte de l'induction magnétique terrestre, une technique plus complexe est nécessaire : nous avons utilisé des tracés de rayon [6] associée à une procédure d'interpolation de données. Pour chaque trajet, on définit deux modes de propagation appelés O et X [7].

Dans le cas où le milieu est considéré comme étant isotrope, nous déterminons donc le nombre de trajets entre un émetteur et un récepteur, alors que si on tient compte de l'induction magnétique terrestre, nous déterminons le nombre de modes de propagation.

### 1.3 Exemples de résultats et conséquences pour la radiogoniométrie

Les figures 2,3,4 présentent respectivement des courbes typiques des angles d'élévation attendus en fonction de la fréquence :

- pour une propagation en considérant l'ionosphère isotrope (le profil utilisé est celui correspondant à la figure 1)
- pour les modes de type O (figure 3)
- pour les modes de type X (figure 4)

Pour simplifier, nous nous sommes limités, pour les figures 3 et 4, au cas où la couche F1 est inexistante. Les calculs sont faits avec  $f_oE = 3$  MHz,  $h_mE = 110$  km,  $y_mE = 10$  km,  $f_oF2 = 9,3$  MHz,  $h_mF2 = 300$  km,  $y_mF2 = 150$  km, la vallée est à 140 km.

La figure 2 montre que pour certaines fréquences, deux ondes peuvent arriver au récepteur pour une réflexion sur une couche donnée : ce sont les ondes dites hautes et basses. L'onde basse est toujours observée alors que l'onde haute n'est présente que pour une plage de fréquences inférieure à la MUF. Ces résultats sont caractéristiques des mesures en sondage oblique [8].

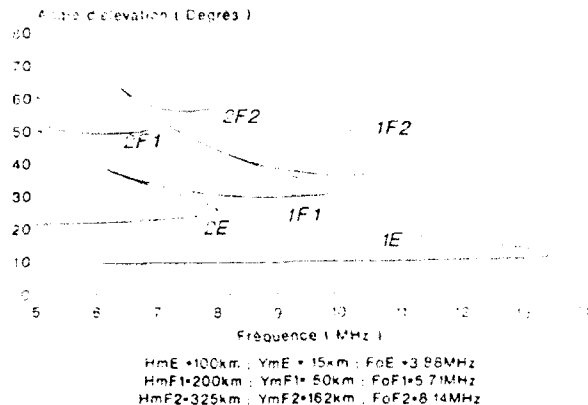


FIGURE 2 - PREVISIONS DES ANGLES D'ELEVATION SUR UNE LIAISON DE 871 KM (JUN 1992)  
ELEVATION ANGLES FORECASTS FOR A 871 KM CIRCUIT LENGTH

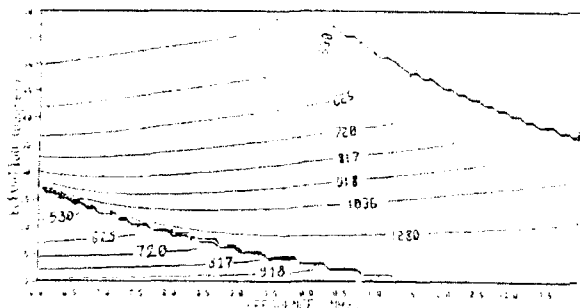


FIGURE 3 - ISOCLINES DE DISTANCE POUR LE MODE O  
SKIP DISTANCE ISOCLINES FOR O MODE

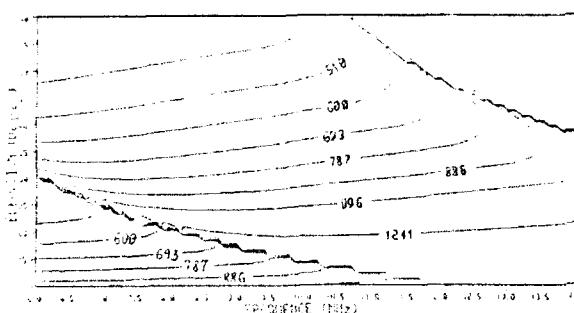


FIGURE 4 - ISOCLINES DE DISTANCE POUR LE MODE X  
SKIP DISTANCE ISOCLINES FOR X MODE

La technique du trace de rayon permet d'estimer l'écart entre les angles d'élévation correspondant aux modes O et X. Dans la zone linéaire de la figure précédente, (zone pour laquelle il n'y a qu'un seul trajet), les valeurs des angles d'arrivée restent très voisines, par contre ces valeurs peuvent notablement différer pour des fréquences voisines de la MUF.

Ceci conduit aux conséquences suivantes pour effectuer une radiogoniométrie :

- plusieurs trajets correspondant à un même mode de propagation peuvent être très voisins. Ceci implique une contrainte de séparation qui ne sera pas toujours aisée à effectuer en particulier avec des réseaux de petite dimension. La résolution spatiale d'un réseau antenne ne permettra donc pas de séparer tous les trajets de propagation.

- deux modes de polarisation correspondant à un même trajet peuvent arriver sous des angles encore plus voisins que dans le cas précédent : compte tenu des limites expérimentales des goniomètres, les techniques qui ne tiennent pas compte des effets de polarisation ne pourront séparer ces modes et les résultats obtenus pourront être erronés.

Enfin, un grand nombre de trajets (et de modes) peuvent être présents sur une liaison : ceci implique un nombre de capteurs suffisants par exemple.

La séparation des signaux obtenus autrement qu'avec un filtrage spatial est donc nécessaire si on veut augmenter les performances des goniomètres. L'établissement d'un modèle de signal réaliste permet de définir des techniques de séparation.

## 2 MODELISATION ANALYTIQUE ET DETERMINISTE DU SIGNAL A LA RECEPTION

### 2.1 Expression générale

Plusieurs approches sont possibles : la première consiste à considérer le canal ionosphérique comme un filtre linéaire variant dans le temps. Le canal est alors décrit soit par sa réponse impulsionnelle  $h(\tau, t)$ , soit par sa fonction de transfert  $H(\omega, t)$ . La détermination analytique de ces deux fonctions n'est pas aisée à partir du moment où le canal est représenté de façon réaliste. De plus, l'approche "filtre linéaire" ne tient pas compte, jusqu'à présent, du caractère vectoriel de la propagation. Ceci signifie en particulier que les fonctions  $h(\tau, t)$  ou  $H(\omega, t)$  incluent le phénomène de couplage champ-antennes. Dans ce qui suit, nous proposons une approche physique du problème déjà abordé par ailleurs [9,10].

En bande étroite, nous écrivons que le signal reçu, sur l'antenne 1, se met sous la forme analytique suivante :

$$s_1 = \sum_{k=1}^N A_k e^{j\theta_k} F_{ik}(G_r E_k A_{2k} P_k)$$

où  $k$  représente le mode incident (O ou X) associé à un trajet.  $N$  : nombre de modes reçus.

$A_k$  est caractéristique de l'amplitude de ce mode et  $\theta_k$  sa phase (en bande étroite  $\theta_k = \omega_k t$ ,  $\omega_k = \omega_0 + \Delta \omega_k(t)$ ,  $\omega_0$  pulsation

du signal émis et  $\Delta \omega_k$  décalage doppler).

$F_{ik}$  est une fonction caractéristique de l'antenne dépendant de son type et lieu d'implantation dans le réseau (paramètre  $G_i$ ) des élévations ( $E_k$ ) et azimut ( $A_{2k}$ ) d'arrivée, et de la polarisation de l'onde incidente ( $P_k$ ) dépendant elle-même de la fréquence, des angles d'arrivée et de la nature du sol au voisinage de l'antenne. Nous examinons le rôle de cette fonction au paragraphe suivant.

Dans le cas où les signaux ont une bande inférieure à 3 KHz, nous écrivons que :

$$s_i(t) = \sum_k m(t - \tau_{gk}) A_k e^{j\theta_k} F_{ik}(G_r, E_k, A_{zk}, P_k)$$

où  $\tau_{gk}$  représente le temps de propagation associé en mode  $k$ , et

$$\theta_k(t) = \int_{t_0}^t \omega_k(t) dt \quad \text{dans toute la bande ; } \Delta \omega_k(t)$$

évalué au temps de phase.

Ceci constitue évidemment une approximation puisque le milieu est dispersif mais qui reste acceptable compte tenu des ordres de grandeur des différents termes.

## 2.2 Effet combiné des antennes et de la polarisation des ondes HF

Connaissant les angles d'arrivée des trajets de propagation, et assimilant la base du profil à l'altitude de sortie des ondes de l'ionosphère, on calcule l'état de polarisation des modes magnétoioniques, en utilisant les conditions aux limites de Budden [7]. Elles donnent en particulier, le rapport ( $R$ ) de polarisation dans le plan d'onde et l'angle  $\alpha$  entre le grand axe de l'ellipse de polarisation et l'horizontale locale. On néglige le terme de collision.

$R$  et  $\alpha$  dépendront de la localisation géographique du récepteur, et de la direction d'arrivée de l'onde. Les composantes du champ électrique associé à chacun des modes sont, par suite, proportionnelles au phaseur  $[1, U(R, \alpha), 0]$ .

Le capteur supposé en diversité de polarisation, est constitué de plusieurs antennes (fouets ou cadres). Il est représenté par une matrice  $T(A_z, E)$ . L'effet de sol  $y$  est inclus.

La connaissance du champ électrique au voisinage des antennes permet de déterminer leur réponse caractérisée par  $F_{ik}$  nombre complexe introduit dans [11]. Il est fonction par suite de l'antenne, de son lieu d'implantation, du type de mode, de la polarisation de l'onde, et, par les conditions de Budden, de la direction d'incidence de l'onde. Il s'écrit :

$$F_{ik} = T(A_{zk}, E_k) \cdot [1, U(R, \alpha), 0] T$$

Une étude complète du comportement des antennes filaires actives a été publiée par ailleurs [11].

Trois figures illustrent les résultats obtenus :

- La figure 5 montre (si l'on admet les conditions précédentes) l'état de la polarisation dans le plan d'onde, des ondes incidentes aux antennes. Sur cette figure, le point milieu représente le point à la verticale de l'observateur. Lorsque le paramètre caractérisant les isocourbes tend vers 1, la polarisation est quasi-circulaire, s'il est égal à zéro, la polarisation est linéaire. Pour Paris, les ondes incidentes sont polarisées circulairement lorsqu'elles arrivent du sud, avec une élévation de l'ordre de  $64^\circ$ . Nous sommes dans ce cas en propagation quasilongitudinale. Des calculs analogues effectués pour d'autres stations [11, 18] montrent la grande variabilité de ces graphes en relation avec le comportement du champ magnétique terrestre.

- Les figures 6 et 7 représentent les modules des réponses à 6 MHz d'antennes filaires actives placées à 5 m de hauteur pour les deux polarisations (O et X) et une station située à Paris. Le sol est caractérisé par sa permittivité ( $\epsilon_r = 15$ ) et sa conductivité ( $\sigma = 10 \cdot 10^{-2}$  S/m). La variation des modules a été interprétée mathématiquement dans [11]. Les différences de phase entre les signaux obtenus, sur les deux antennes croisées, sont également données dans [18]. Les antennes cadres très souvent utilisées en radiogoniométrie HF ont des comportements semblables à ceux des antennes filaires placées dans les mêmes conditions expérimentales.

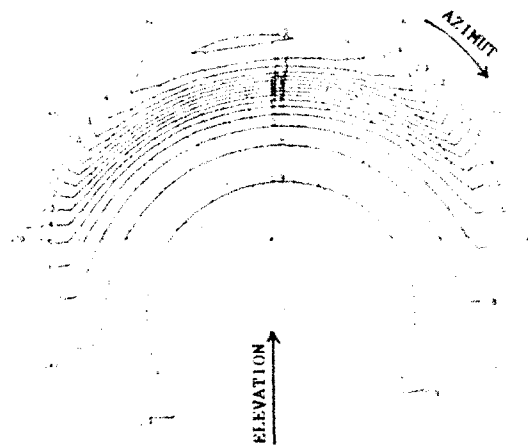


FIGURE 5 - ABAQUE DE POLARISATION (PARIS-6 MHz)  
POLARIZATION CHART (PARIS-6 MHz)

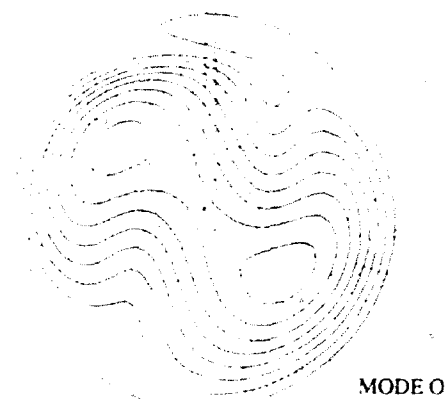


FIGURE 6

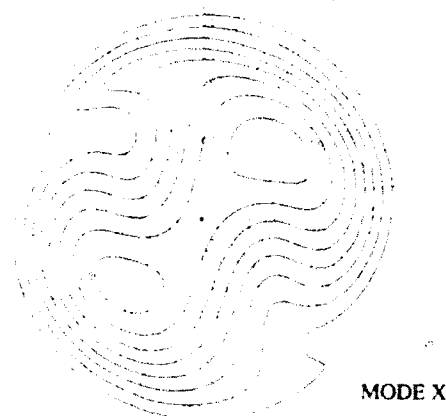


FIGURE 7

AMPLITUDE DES SIGNAUX, ANTENNE, E-W  
SIGNAL INTENSITY, E-W ANTENNA

### 2.3 Décalage doppler

Nous considérons des mouvements lents de l'ionosphère, ou d'une couche particulière, tels que par exemple le mouvement ascensionnel des couches ionosphériques au cours de la journée, et le mouvement pseudopériodique induit par le passage d'une onde atmosphérique de gravité.

Si  $v(r,t)$  représente la vitesse de ce mouvement, l'indice s'écrit de façon générale, sous la forme :

$$n(r,t) = n(r - v(r,t), t)$$

Pour des mouvements suffisamment lents et de petites échelles vis à vis de la distance de propagation, l'allure des variations donnée à  $v(r,t)$  se retrouve sur celles du temps de phase et de groupe. En notant  $f_0$  la fréquence de la porteuse, la dérivée du chemin de phase donne l'allure du doppler attaché au  $k$ ème mode de propagation particulier. On écrit :

$$f_{dk}(t) = -f_0 \cdot d(P_{dk}(t,f)) / dt$$

$P_{dk}(t,f)$  est le temps de phase du mode  $k$ .

Pour simplifier la génération du signal pour une période de simulation typiquement inférieure à 1 heure, nous donnons au doppler une allure simple de type sinusoïdale déjà introduite dans [9 et 15]. Les valeurs d'amplitude des oscillations et leur période sont déduites d'expérimentations [10,12].

Ainsi l'amplitude des mouvements ionosphériques (en particulier celle due aux ondes atmosphériques de gravité) croît jusqu'à une altitude voisine du maximum d'ionisation [13]. Nous écrivons que :

$$\omega_k \cdot t = \omega_0 \cdot t + a \frac{E_k(d)}{100} \cdot T \cos\left(\frac{2\pi t}{T} + \Phi_k\right)$$

où :

$E_k$  : est l'angle d'élévation exprimé en degrés

$T$  : est la période d'une perturbation ( $T$  typique associée à une onde de gravité voisine de 15 mn)

$\omega_0$  : est la pulsation de la porteuse (nulle en bande de base)

$\Phi_k$  : est la phase à l'origine

$a$  : terme d'amplitude associé à la distance émetteur-récepteur.

Ce dernier formalisme néglige la composante continue du doppler, mais rend bien compte d'observations réalisées par ailleurs [8,10,12]. Le décalage doppler est d'autant plus important que l'élévation est élevée, ce qui, pour une liaison donnée, traduit l'influence de la pénétration dans le milieu ionosphérique. Typiquement, pour des distances de 1000 km, on prendra  $a = 1$ , ce qui conduit à des effets dopplers ionosphériques de l'ordre de 0,9 Hz au maximum.

Nous avons vu que pour un même trajet, les trajectoires des modes O et X restent très proches ; les dopplers correspondants seront très voisins et il sera difficile de séparer ces ondes en des temps de l'ordre de quelques secondes. Seule une technique utilisant la diversité de polarisation permettra de le faire.

La figure 8 présente les signaux observés sur 2 cadres croisés dans le cas d'une liaison de 620 km. On constate un comportement très différent, ce qui est conforme au modèle développé.

### 3 GONIOMETRIE : PROBLEME DIRECT

Nous considérons dans cette partie le cas de réseaux de petites bases avec lesquels on souhaite estimer les angles d'arrivée des multitrajets ionosphériques. Nous présentons des résultats de simulation et nous n'envisageons pour cela que les effets de propagation modélisés dans la partie 2 : nous supposons également que le système antenne-électronique-traitement, est parfait, et que les effets de bruit ou de couplage sont négligés.

Les simulations montrent notamment comment les multitrajets affectent les procédures dites classiques et dans quelles limites un traitement de données permet de les améliorer.

Nous présentons ensuite une première approche de goniométrie par une technique de traitement de signal simple : elle consiste à séparer les multimodes par analyse spectrale [14].

L'attention est portée davantage sur l'application récente de l'algorithme Haute Résolution MUSIC à un système HF expérimental. Nombreuses références sont faites à Rogier et al [15], en ce qui concerne l'implémentation de l'algorithme, la description du système et sa validation expérimentale.

Nous montrons également l'intérêt d'ajouter un filtrage de polarisation (ce qui revient à travailler avec un réseau d'antennes en diversité de polarisation) dans le cas d'une trop forte corrélation entre modes de propagation.

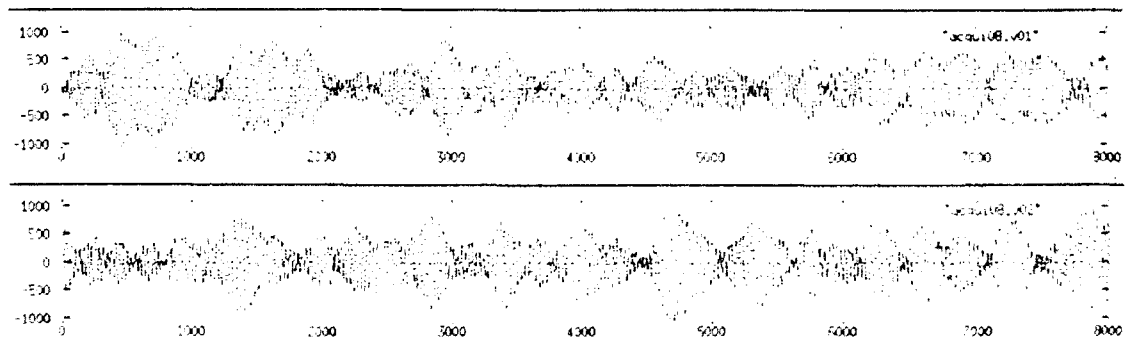


FIGURE 8 - EXEMPLE DE SIGNAUX REÇUS PENDANT 80 SECONDES SUR DEUX CADRES CROISÉS : 6.09 MHz  
LIAISON LUXEMBOURG - RENNES  
80 SECONDS OF SAMPLES RECEIVED ON TWO CROSSED-LOOP ANTENNAS : 6.09 MHz  
CIRCUIT LUXEMBOURG - RENNES



### 3.1 Comportement des interféromètres

Plusieurs types de radiogoniomètres ont été développés en particulier pour la localisation en VHF et UHF. Les techniques utilisées ont souvent été transférées pour réaliser des goniomètres dans la gamme HF (interféromètres, goniomètres doppler, balayage de lobe). Le paragraphe précédent nous a montré la complexité des signaux reçus. Pour illustrer leur comportement en HF, nous nous limitons à la présentation du comportement de l'interféromètre, les autres techniques conduisent à des résultats analogues.

#### *Goniométrie par interférométrie instantanée avec utilisation de capteurs composés d'une seule antenne*

Les angles d'arrivée sont déterminés à partir des différences de phase des signaux observés sur chacune des antennes. Aucune amélioration dans le traitement n'est apportée et c'est pourquoi nous qualifions ces interféromètres de "classiques".

L'application aux goniomètres de dimensions faibles par rapport à la longueur d'onde, du modèle et des relations de l'interférométrie conduisent aux conclusions suivantes :

a) en l'absence d'effet doppler différentiel et pour des amplitudes de signaux identiques pour chacun des modes de propagation, le système calcule une élévation "moyenne" telle que :

$$\cos E_{\text{moyen}} = \frac{1}{N} \sum_{k=1}^N \cos E_k$$

N : nombre d'ondes incidentes au système.

b) La présence d'un doppler différentiel intermode introduit des variations importantes de l'angle moyen calculé autour de la valeur précédente.

Dans tous les cas l'azimut pour une même liaison émetteur-récepteur est calculé correctement dans la mesure où on suppose que les ondes incidentes ont le même azimut.

La figure 9 présente le résultat d'une goniométrie simulée sur ce principe et réalisée en supposant 3 modes incidents ( $E_k = 30^\circ, 60^\circ, 62^\circ$ ) de même azimut ( $0^\circ$ ) et d'amplitudes égales et incluant un doppler différentiel. On peut constater la variation des angles calculés. Ces formes de fluctuations sont observées fréquemment et sont explicables avec le formalisme proposé.

#### *Goniométrie par interférométrie instantanée avec utilisation de capteurs à plusieurs antennes*

L'emploi de plusieurs antennes pour constituer un capteur permet de réaliser des filtrages de polarisation tels qu'ils ont été définis de façon générale par Compton [17], et en HF par Bertel et al [16] et Rojas [10]. De façon industrielle, en HF, la solution souvent retenue consiste à sommer les signaux obtenus sur 2 antennes à polarisation horizontale disposées perpendiculairement en déphasant l'un d'eux de  $\pm \pi/2$ .

Ceci permettrait d'éliminer l'une des ondes O ou X d'un même trajet si les signaux correspondants étaient toujours déphasés de  $\pm \pi/2$ . Ce n'est pas toujours le cas [18].

Cette procédure dans le cas d'un trajet peut améliorer la goniométrie comme le montre la figure 10.

Dans le cas où il y a plusieurs trajets, le problème se ramène au cas précédent (figure 11).

Ces remarques montrent que le comportement de ce type de radiogoniomètre n'est pas satisfaisant et que l'utilisation des résultats obtenus doit être associée à une analyse à postériori.

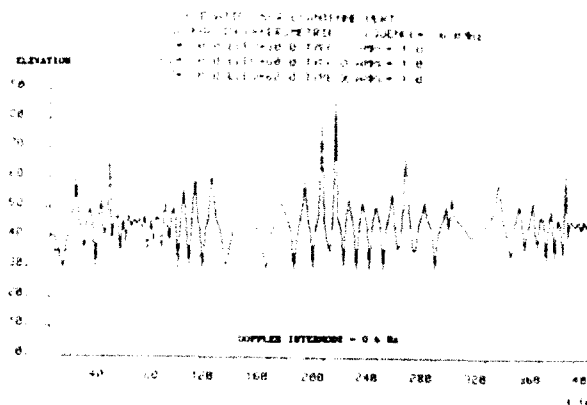


FIGURE 9 - SIMULATION DE RESULTATS D'UNE INTERFEROMETRIE CLASSIQUE  
SIMULATION RESULTS OF CLASSICAL INTERFEROMETRY BASED SYSTEM

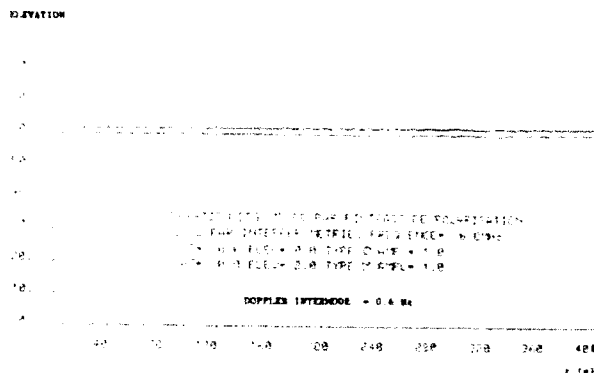


FIGURE 10 - SIMULATION DE RESULTATS D'UNE INTERFEROMETRIE AVEC FILTRAGE DE POLARISATION (2 MODES O ET X INCIDENTS)  
POLARIZATION FILTERING APPLIED TO A CLASSICAL INTERFEROMETRY BASED SYSTEM (2 INCIDENT MODES O AND X)

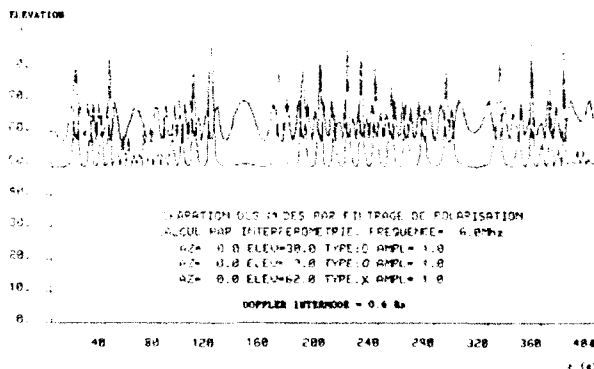


FIGURE 11 - SIMULATION DE RESULTATS D'UNE INTERFEROMETRIE AVEC FILTRAGE DE POLARISATION (3 MODES O, O ET X INCIDENTS)  
POLARIZATION FILTERING APPLIED TO A CLASSICAL INTERFEROMETRY BASED SYSTEM (3 INCIDENT MODES O, O AND X)

### Amélioration de la technique précédente et interférométrie à traitement de données

Le système Skyloc [19] représente une amélioration de la technique précédente. Elle consiste à conserver les données qui vérifient une certaine linéarité du plan d'onde le long du réseau. Ces données particulières vérifient l'hypothèse de propagation unimodale (QUMP). Le cumul des résultats d'interférométrie sous forme d'histogrammes permet un gain sensible dans la précision et l'exploitation des goniométries. La distinction entre modes O et X est cependant difficile.

Les deux techniques précédentes ne prennent pas en compte de modèle de propagation ni de signal HF. Les résultats sont bien sûr corrects si l'hypothèse QUMP est respectée. En pratique cependant la probabilité d'avoir ces conditions diminue rapidement dès que le nombre de trajets dépasse 2 [20].

L'exploitation des résultats en terme de goniométrie repose ensuite sur la variabilité du canal HF : en considérant  $N$  ondes incidentes,  $N-1$  s'annulent par interférence. Reste un mode qui apparaît distinctement et peut être goniométré, aucun effort d'identification n'est fait à ce stade. On espère ensuite que les  $N$  ondes vont se combiner entre elles et apparaître à tour de rôle.

L'avantage est dans la simplicité, le désavantage dans le temps d'établissement d'histogramme : typiquement 10 à 20 mn d'après les résultats publiés [21].

### 3.2 Goniométrie utilisant la technique du balayage de lobe et traitement des données

Il avait été montré que l'utilisation d'un réseau phasé de grande dimension [22] associé à la technique de "beamforming" ne permettait pas avec un seul balayage d'obtenir les valeurs des angles d'arrivée. Le réseau, compte tenu de sa dimension peut permettre la séparation des trajets mais les antennes utilisées (verticales) ne permettent pas la séparation des modes O et X d'un même trajet.

Une technique de moyennage basée sur la remarque présentée au paragraphe 3.1 conduit, sous certaines contraintes, à une bonne détermination des angles d'arrivée relatifs aux trajets. Nous ne développons pas ici ce sujet traité dans un autre article [23] de cette même conférence.

### 3.3 Association d'une technique d'interférométrie et d'analyse spectrale

Dans [14], nous avons envisagé une première approche pour séparer les multimodes : nous travaillons sur porteuse et nous séparons les différents modes par analyse spectrale linéaire (FFT) ou non linéaire (Burg-Marple) super résolutive complétée ou non par un filtrage de polarisation.

Si la résolution est suffisante, on déduit les angles d'arrivée des différentes modes, des phases associées aux maxima du spectre. Le cas contraire permet d'accéder au mieux, aux angles des trajets de propagation.

La procédure est inadéquate aux instants où les dopplers intermodes sont nuls. La cohérence des résultats de goniométrie au cours du temps, ou le cumul de ceux-ci en histogramme, permet d'éliminer les ambiguïtés dues à ces mauvaises goniométries.

Le temps de stationnarité du signal limite l'emploi de la FFT à des durées de mesure de 20 à 50 secondes environ, selon le nombre de modes présents. Les méthodes de Burg et de Marple apporte si nécessaire, une meilleure résolution que la précédente; à résolution égale, et pour un rapport signal sur bruit suffisant, elles permettent également de diminuer le temps nécessaire pour goniométrer un émetteur. La figure 11 présente un résultat obtenu avec les mêmes données que pour les figures 9 et 10. Le doppler intermode est supposé être de 0.1 Hz et la durée de l'échantillonnage de 12.7 s.

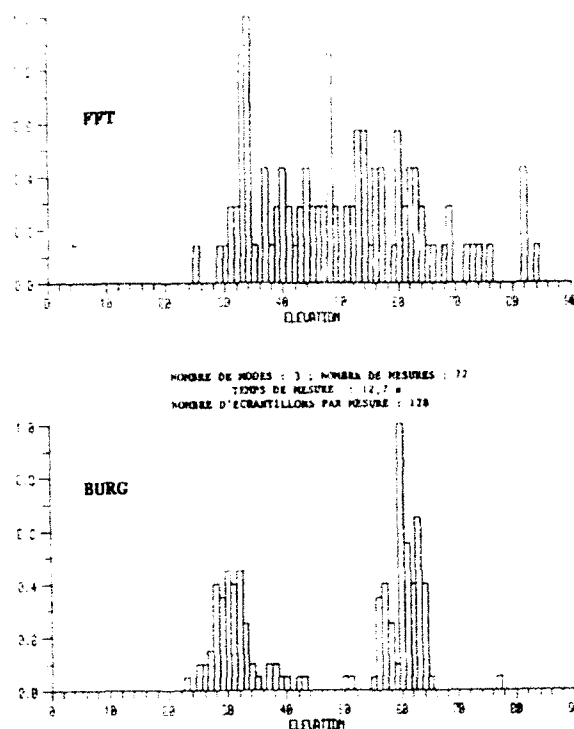


FIGURE 12 - INTERFEROMETRIE PAR ANALYSE SPECTRALE  
SPECTRAL ANALYSIS INTERFEROMETRY METHOD

### 3.4 Exemple d'un algorithme Haute Résolution appliqué à la HF : MUSIC

MUSIC : Multiple Signal Classification, est une des méthodes de goniométrie basée sur les propriétés de décomposition en éléments propres de la matrice de covariance. Ces méthodes ont été dans la dernière décennie l'objet de nombreuses recherches et s'appliquent indifféremment à l'analyse spectrale et à la goniométrie. En HF, citons parmi d'autres, Gething [20] qui utilise ces propriétés pour des antennes adaptatives. Johnson [24] qui rapporte la simulation de MUSIC à l'estimation de l'azimut d'ondes de sol uniquement.

Trois difficultés à l'application de cette technique : la corrélation qui, si elle est totale, détruit le pouvoir haute résolution de l'algorithme ; l'identification du nombre de modes présents ; la connaissance en gain et phase du réseau antennaire [25].

Ces différents points sont particulièrement sensibles en HF. La partie 2 a permis de proposer une modélisation pour deux d'entre eux : la corrélation, et la réponse des antennes.

A l'exemple de [24], nous supposons le nombre de modes fixe, et nous ne montrons que l'effet de la corrélation et du filtrage de polarisation sur la goniométrie. Par suite, nous utilisons uniquement la méthode de moyennage direct pour l'estimation de la matrice de covariance.

Une étude plus générale de cette technique est présentée par G. Multedo dans un autre article de cette conférence [26].

#### Présentation

L'application de MUSIC nécessite une formulation matricielle du signal reçu sur le réseau. Le vecteur d'observation  $x(t)$  formé des signaux  $s(t)$  sur les différents capteurs, s'écrit :

$$x(t) = A.s(t) + b(t)$$

$x(t)$  a la dimension du nombre de récepteurs (P) constituant le réseau  
 $s(t)$  de dimension N, représente les modes incidents, dont l'expression est donnée dans la partie 2  
 $b(t)$  est le bruit additif sur chacun des capteurs du réseau.

A une matrice dont chaque colonne a ( $Az_k, E_k$ ) représente un modèle de plan d'onde incident.

L'algorithme MUSIC utilise les vecteurs propres de la matrice de covariance spatiale  $R_{xx}$ , ou de son estimée obtenue par moyennage sur une durée  $T_a$  de la matrice  $x(t).x(t)^+$ . En supposant N modes présents, non totalement corrélés, les vecteurs propres correspondants aux N plus grandes valeurs propres décrivent l'espace signal. Les P-N vecteurs propres restant, décrivent l'espace bruit.

L'hermiticité de  $R_{xx}$  et par suite, l'orthogonalité de ces 2 espaces, permet de définir une fonction qui est maximum pour les directions d'arrivée des ondes incidentes. Cette fonction porte indifféremment le nom de spectre de MUSIC dans [27,28,29] ou de fonction du goniomètre dans [30]. Elle s'écrit :

$$f(Az, E) = \frac{a^+(Az, E) a(Az, E)}{\sum_{k=N+1}^P |V_k^+ a^+(Az, E)|^2}$$

où  $V_k$  représentent les vecteurs propres.

Le nombre de modes pouvant être obtenu par MUSIC est théoriquement égal au nombre de capteurs du réseau moins un. De plus, dans le cas où la matrice de covariance est bien estimée, ainsi que la réponse du réseau d'antennes, la résolution est infinie. En pratique, les hypothèses qui permettent d'atteindre ces performances ne sont pas vérifiées.

#### Allure de la corrélation entre deux modes

Le modèle de signal du paragraphe 2 avec son doppler lentement variable, rend celui-ci non stationnaire sur un intervalle de temps  $T_a$ .

Pour le cas de deux modes de propagation, la matrice de covariance  $R_{xx}(t_0, T_a)$  estimée par moyennage direct, fait intervenir le terme  $r(t_0, T_a)$ , qui selon le modèle de signal, est le produit de la fonction d'autocorrélation  $\Gamma_m(\tau)$  de la modulation  $m(t)$  et du terme de corrélation introduit par le doppler différentiel. Ce dernier dépend de la variable  $\Delta\omega$ . Nous écrivons, comme Rogier et al. [15] :

$$r(t_0, T_a) = \Gamma_m(\tau_{g1} - \tau_{g2}) \cdot \rho(t_0, T_a)$$

$$\rho(t_0, T_a) = \frac{1}{T_a} \int_{t_0}^{t_0 + T_a} \exp \left[ j \Delta \omega_{12} \sin \left( \frac{2\pi}{T} \right) \right] dt$$

$\omega_{12}$  : amplitude du doppler entre les modes 1 et 2 de temps de groupe  $\tau_{g1}$  et  $\tau_{g2}$

T : période de l'onde atmosphérique de gravité  
 $t_0$  : temps initial de la mesure.

La figure 13 représente les isoclines du module de la corrélation intermode;  $m(t)$  est un bruit blanc de bande B pour simplifier;  $t_0$  est pris nul. La différence de temps de propagation entre les modes (axe horizontal) ou une intégration sur le temps (axe vertical), permet de diminuer la corrélation.

Nous nous plaçons par la suite dans le cas défavorable d'une porteuse pure, ou seule une intégration sur le temps permet la décorrélation. La figure 14 montre cet effet, pour différentes amplitudes de doppler intermode, en fonction de l'instant initial de mesure  $t_0$ .

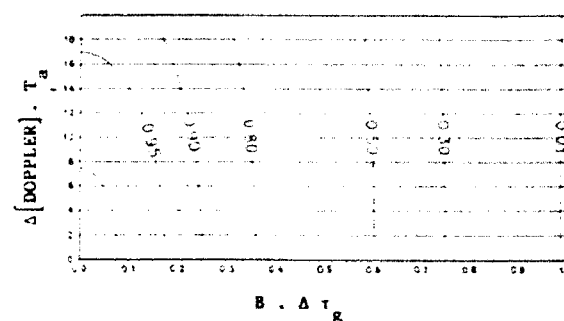


FIGURE 13 - ISOCLINES D'AMPLITUDE DE LA CORRELATION ENTRE 2 MODES DE PROPAGATION  
 CORRELATION AMPLITUDE ISOCLINES BETWEEN 2 MODES OF PROPAGATION

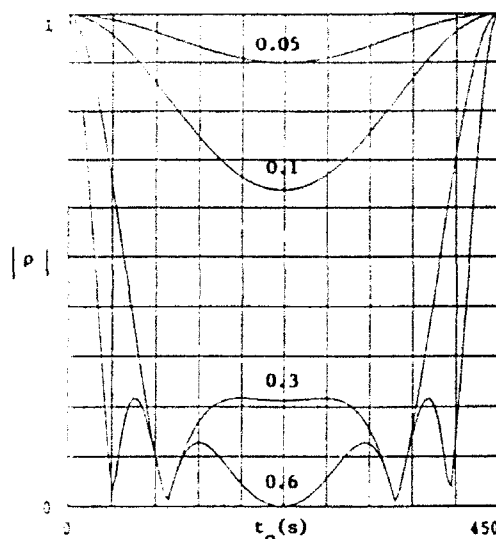


FIGURE 14 - AMPLITUDE DE LA CORRELATION EN FONCTION DE  $t_0$  POUR DIFFERENTES VALEURS DU DOPPLER, TEMPS D'INTEGRATION : 5 s.  
 CORRELATION AMPLITUDE VERSUS  $t_0$  FOR DIFFERENT DOPPLER SHIFT AMPLITUDES, INTEGRATION TIME : 5 s

#### Influence de la corrélation sur la goniométrie

Nous considérons le cas simple d'un réseau linéaire de 3 capteurs, et de 2 modes incidents. Nous avons choisi un mode à 30 degrés, le second à 60 degrés d'élévation et de même azimut. Le signal est une porteuse pure, la matrice de covariance est estimée par moyennage direct, et nous utilisons les deux valeurs propres les plus importantes, sans discrimination sur leurs valeurs.

Les figures 15 et 16 montrent en parallèle l'évolution de la corrélation intermode, le comportement des deux valeurs propres "signal" et les résultats de goniométrie, ceci en fonction de l'instant initial  $t_0$  de mesure et pour une durée d'intégration  $T_a$  non optimisée par rapport à [15]. Dans ces figures, à  $t_0 = 0$ , les signaux sont corrélés, la période T est prise égale à 900 secondes.

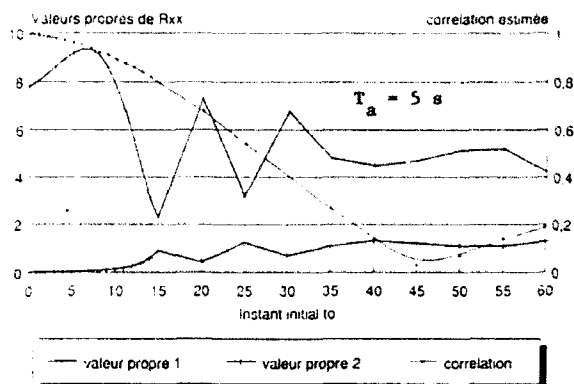


FIGURE 15 - EFFET DE LA CORRELATION SUR LA DETERMINATION DES VALEURS PROPRES DE LA MATRICE DE COVARIANCE  
EFFECT OF THE CORRELATION LEVEL ON THE BEHAVIOUR OF THE TWO MAIN EIGEN VALUES OF THE COVARIANCE MATRIX

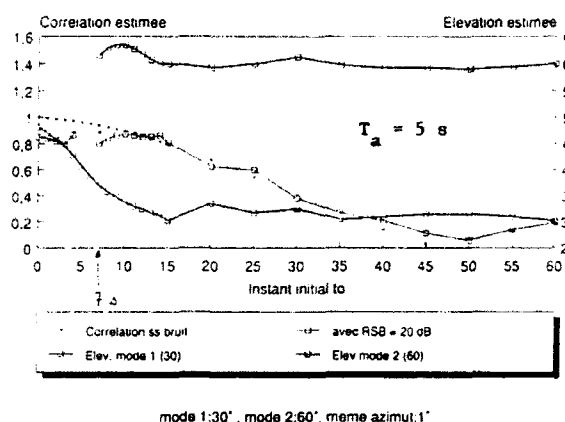


FIGURE 16 - ESTIMATION DES ELEVATIONS DE 2 MODES (30°, 60°) ET AMPLITUDE DE LA CORRELATION  
ESTIMATED ELEVATION ANGLES FOR 2 MODES (30°, 60°) AND CORRELATION AMPLITUDE

Pour une corrélation inférieure à 95%, l'algorithme détecte les deux modes présents. La goniométrie est considérée comme correcte pour une valeur inférieure à 85%.

#### Influence du filtrage de polarisation

Les capteurs du réseau sont constitués cette fois de 2 antennes foyers horizontaux, orientées Est-Ouest et Nord-Sud. Le filtrage de polarisation consiste à sommer les signaux reçus sur ces deux antennes, après un déphasage de  $\pm \pi/2$  de l'un d'eux. Le capteur obtenu est une antenne à polarisation circulaire droite ou gauche, qui permet d'atténuer l'un des modes O, X des trajets de propagation.

La théorie magnétoionique et les abaques, montrent cependant que ce filtrage n'est pas optimum dans toutes les directions. Les résultats suivants le confirment également.

Les simulations prennent en compte les deux trajets précédents avec les deux modes magnétoioniques présents pour chacun d'eux: 30, 32, 60, 62 sont les angles d'élévation attendus. L'effet de sol n'est pas pris en compte

Nous présentons deux cas selon que les modes sont corrélés à 98% (figure 17) ou complètement décorrélés (figure 18). Les résultats sont fonction de l'azimut des multitrajets.

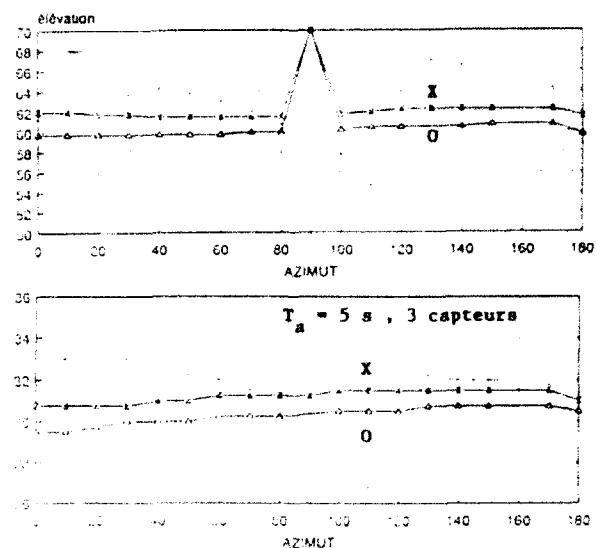


FIGURE 17 - ESTIMATION DES ELEVATIONS PAR MUSIC AVEC FILTRAGE DE POLARISATION, EN POINTILLES : RESULTATS SANS FILTRAGE, SIGNAUX CORRELÉS  
ESTIMATED ELEVATION ANGLES BY MUSIC WITH POLARISATION FILTERING, IN DOT LINES : RESULTS WITHOUT FILTERING, CORRELATED SIGNALS

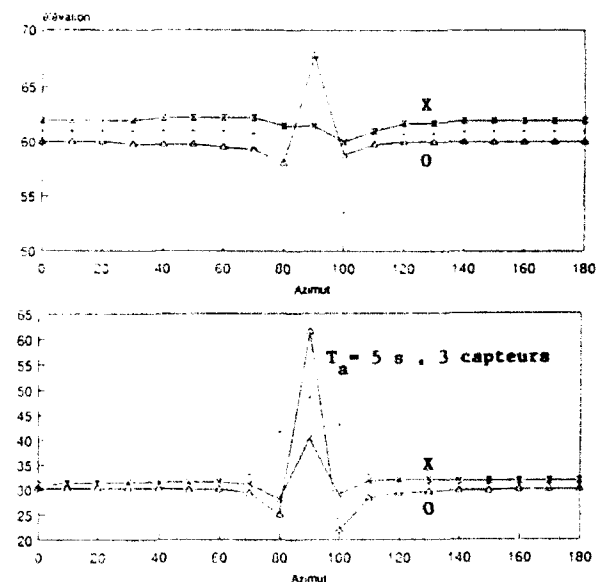


FIGURE 18 - ESTIMATION DES ELEVATIONS PAR MUSIC AVEC FILTRAGE DE POLARISATION, EN POINTILLES : RESULTATS SANS FILTRAGE, SIGNAUX DECORRELÉS  
ESTIMATED ELEVATION ANGLES BY MUSIC WITH POLARISATION FILTERING, IN DOT LINES : RESULTS WITHOUT FILTERING, DECORRELATED SIGNALS

Sans filtrage, la corrélation intermode perturbe la goniométrie. L'erreur varie avec l'azimut, c'est-à-dire avec la polarisation de l'onde. Sans corrélation, on obtient un résultat intéressant dans le sens où l'algorithme estime la valeur moyenne des 2 modes (de même énergie ici). Dans les deux cas, pour un azimut de 90°, le filtrage devient inefficace. Ceci est aisément interprétable quand on tient compte de la réponse des antennes.

La procédure avec filtrage sépare les deux modes. l'erreur reste inférieure au degré. Les courbes montrent une faible variation avec l'azimut et le niveau de corrélation.

#### 4 LOCALISATION : PROBLEME INVERSE

Dans ce qui suit, nous ne présentons que quelques aspects du problème inverse. Nous supposons que le problème direct a été bien traité et que les angles d'arrivée correspondant aux ondes les plus "énergétiques" sont bien déterminés.

##### 4.1 Influence d'une incertitude sur le profil de densité électronique

Pour un cas typique (distance D émetteur-récepteur voisine de 1300 km, milieu de journée juin 1992), nous présentons la variation de la distance émetteur-récepteur pour un angle d'élévation donné (30°), en fonction de la variation de  $h_m F_2$  (figure 20), et de  $f_o F_2$  (figure 19). Nous constatons qu'une erreur de  $\pm 0.5$  MHz autour de la valeur  $f_o F_2$  prédite se traduit par une variation relative de distance  $\Delta D/D$  pour le cas présenté de 17 %. Une variation de 50 km sur  $h_m F_2$  conduit de la même façon à un  $\Delta D/D$  de 30 %.

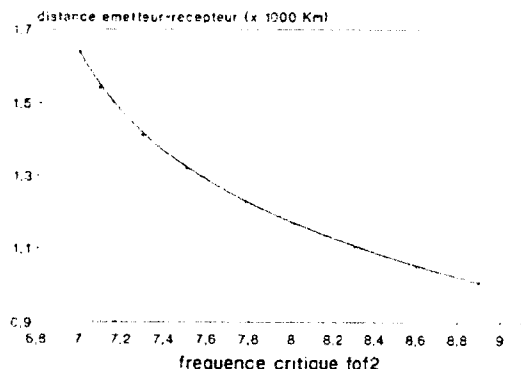


FIGURE 19 - VARIATION DE LA DISTANCE EMETTEUR-RECEPTEUR EN FONCTION DE  $f_o F_2$   
SKIP DISTANCE VARIATION VERSUS  $f_o F_2$

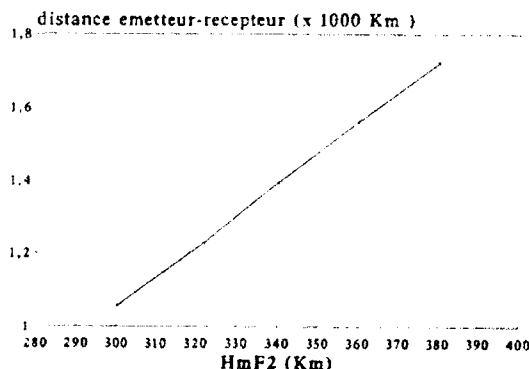


FIGURE 20 - VARIATION DE LA DISTANCE EMETTEUR-RECEPTEUR EN FONCTION DE  $h_m F_2$   
SKIP DISTANCE VARIATION VERSUS  $h_m F_2$

Les résultats montrent l'importance que revêt la connaissance du profil de densité. Afin de disposer d'informations sur celui-ci, plusieurs méthodes peuvent être utilisées. Nous les décrivons succinctement.

##### Utilisation d'un sondeur vertical au point de réception

Cette technique est employée par différentes équipes. Elle nécessite l'implantation d'un système "actif" au point de réception. Son avantage réside dans l'utilisation possible d'une méthode de détermination de la distance basée sur la représentation (MQP) du profil [31]. Cette méthode doit toutefois être combinée avec une transposition du profil de densité en fonction du point de réflexion.

##### Procédure de réactualisation du profil de densité

Cette méthode a été décrite par ailleurs [5], elle s'appuie sur :

- des données de propagation issues de prévisions
- des mesures effectuées par le goniomètre sur des émetteurs connus fonctionnant dans une zone voisine de celle où le système à localiser se trouve. A partir de ces mesures, une réactualisation partielle du profil de densité est effectuée. Les premiers résultats obtenus sont très encourageants.

Cette méthode présente l'avantage de ne pas nécessiter de système actif (émetteur), par contre, elle suppose la connaissance de la localisation d'autres émetteurs dans des zones voisines de l'émetteur à localiser.

##### 4.2 Influence de l'inclinaison des couches ionosphériques

L'inclinaison des couches ionosphériques (tilts) est liée à la fois à la variation diurne du milieu et aux perturbations de petites ou moyennes échelles associées aux ondes atmosphériques de gravité ou orages ionosphériques. Les couches réfléchissantes peuvent alors schématiquement revêtir différentes formes : plan incliné, ondulations, surfaces concaves ou convexes. Ceci conduit :

- à faire varier les angles d'arrivée (azimut et élévation)
- à augmenter parfois le nombre de trajets et modes.

De nombreuses études ont déjà été publiées sur ce sujet [32]. Nous avons vu qu'il était difficile de séparer les signaux correspondant à chacun des modes de propagation. En conséquence, les résultats publiés sur les variations d'angles d'arrivées, doivent être utilisés avec précaution.

La figure 21 donne un exemple de variation de l'azimut d'arrivée avec l'élévation pour un trajet donné de 352 km (Cholet-Coulommiers). Le dispositif expérimental utilisé est celui décrit par Rogier et al [15]. Les mesures ont été effectuées sur 4 jours en septembre 1991 au moment d'une activité magnétique très variable. Elles conduisent à des angles d'élévation variant de 20° (propagation en 1 bond par réflexion sur la couche E), à 75° (propagation en 2 bonds par réflexion sur la couche F2).

L'influence de la variation diurne du milieu ionosphérique apparaît clairement. On constate en effet que :

- aux faibles élévations, azimut estimé et azimut géométrique [236°] diffèrent de  $\pm 1^\circ$  environ, ce qui correspond à la précision attendue sur une mesure perturbée par les mouvements de petites échelles.
- aux élévations plus élevées, l'écart entre les valeurs expérimentale et géométrique (pour l'azimut) s'amplifie. Ceci peut se concevoir car la précision sur la détermination de l'azimut décroît avec l'élévation. Toutefois, cet écart n'est pas distribué de façon aléatoire, il apparaît comme étant : négatif le matin, nul vers midi, positif le soir.

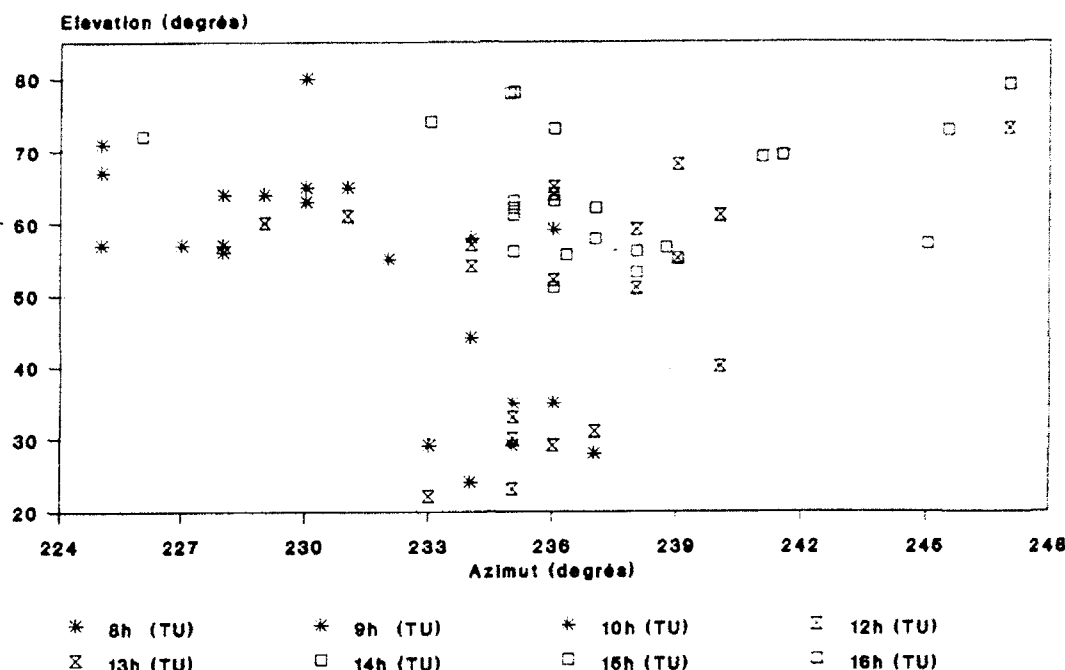


FIGURE 21 - EFFET DES TILTS : ELEVATION EN FONCTION DE L'AZIMUT  
TILTS EFFECTS : ELEVATION VERSUS AZIMUTH

On met donc en évidence un phénomène qui est de nature géophysique et non instrumentale. La liaison est Sud-Ouest → Nord-Est. Le signe des écarts peut s'interpréter en termes d'inclinaison des couches de la région F : le matin elles sont plus basses à l'est qu'à l'ouest, et au sud qu'au nord. L'altitude de la couche E varie peu. Une liaison s'appuyant sur une réflexion spéculaire sur de telles couches, conduit à déterminer un azimut inférieur à l'azimut local géométrique le matin (Pour ce cas TU = TL). Le soir, le phénomène est inversé. A midi, les couches sont supposées quasiment horizontales et l'écart est statistiquement nul.

#### 4.3 Conséquences expérimentales

Les deux paragraphes précédents montrent que pour être performante, la localisation HF nécessite une connaissance en 4 dimensions (altitude, latitude, longitude, temps) du milieu ionosphérique. Les techniques de réactualisation elles-mêmes doivent tenir compte de cette variabilité.

Du point de vue opérationnel, lorsqu'on dispose de plusieurs angles d'arrivée, la variabilité du milieu conduit d'abord à utiliser les angles correspondant à une propagation par la couche E relativement stable pour déterminer la distance. La discrimination entre une liaison par un ou plusieurs bonds, peut alors être réalisée en employant les valeurs des autres angles d'arrivée.

### 5 CONCLUSION

Nous avons abordé en parallèle dans cet article, les deux problèmes propres à la radiolocalisation HF. Nous avons montré comment certains phénomènes de propagation pouvaient influencer chacun d'eux.

Le problème direct a montré les nombreux multitrajets qu'un système de goniométrie a à résoudre. L'anisotropie attachée à ceux-ci un mode O et X d'incidences très voisines. La modélisation MQP permet dans tous les cas d'obtenir rapidement les différents paramètres de propagation (élévation, phase, groupe, doppler) que prend en compte la simulation du signal reçu.

Le modèle de signal développé a permis d'interpréter le comportement de certains systèmes de goniométrie, et comment un filtrage de polarisation industriel conduit à recouvrer une partie des capacités perdues. On peut également conclure que la conception de ces systèmes a longtemps souffert de l'image trop aléatoire que l'on avait du canal ionosphérique.

Le problème inverse a montré le rôle important que revêt la connaissance du profil de densité si on veut effectuer une localisation avec précision. Il a souligné également l'effet de l'inclinaison des couches et la nécessité de leur prise en considération pour améliorer les performances des systèmes.

Enfin, les travaux sur le filtrage de polarisation abordés ici, montrent le potentiel que l'on pourrait en attendre en HF. Cette extension nécessite en contrepartie un effort important de modélisation des antennes.

La notion de haute résolution et notamment le nombre maximum de sources pouvant être résolues par ce type d'algorithme, est délicate à vérifier expérimentalement. Cette observation n'est pas spécifique de la HF. La mise en œuvre d'une telle technique nécessite une maîtrise parfaite de l'électronique et des méthodes de calibration du réseau antennaire.

### 6 REMERCIEMENTS

Ce travail a été soutenu par les contrats DRET (n° 89.34.062 et n° 90.473), THOMSON (n° 007T.89) et CNET (n° 89.8B.068). Les auteurs tiennent également à remercier Messieurs Multédo et Rogier de la Société Thomson-CSF, division RGS, pour les nombreuses discussions sur le sujet, et les données expérimentales qui ont permis cette étude.

## 7 REFERENCES

- 1 Hill, J.R., "Exact ray paths in a multisegment quasiparabolic ionosphere", *Radio Science*, vol. 14, n° 5, sept-oct. 1979, pp 855-861.
- 2 Bennett, J.A., Dyson, P.L., "A model of the vertical distribution of the electron concentration in the ionosphere and its application to oblique propagation studies", *J. Atmos. Ter. Phys.*, vol. 50, 1988, pp 251-262.
- 3 Bakert, D., Lambert, S., "Range estimation for SSL HFDF systems by means of a MQP ionospheric model", *IEE Proceedings*, 136, part H 120.
- 4 Titheridge, J.E., "Ionogram analysis with the generalised program POLAN", *World Data Center a for Solar Terrestrial Physic Report UAG-93*, Boulder, Colorado, 1985.
- 5 Bertel, L., Brousseau, C., Baltazart, V., "Prospectives de réactualisation de profils de densité électronique ionosphérique", *Journées d'études propagation électromagnétique dans l'atmosphère du décimétrique à l'angström*, Rennes, 9 et 10 octobre 1991.
- 6 Jones, R.M., Stephenson, J.M., "A versatile three dimensional raytracing computer program for radio waves in the ionosphere, OT Report 75-76, Off. of Telecom., Boulder, Colorado, 1975.
- 7 Budden, K.G., "The propagation of radio waves", 1988, édition Cambridge Press.
- 8 Le Roux, Y., Lassudrie-Duchesne, P., Fleury, R., Jolivet, J.P., Menard, J., "Measurement capabilities of an HF channel evaluation system", *IEE Proceedings* 339, juillet 1991.
- 9 Gourvez, P., Bertel, L., Rojas-Varela, J., "Measurements and analysis of the HF field received with a tripole antenna", *IEE Conf. Pub.*, n° 274, 1987, pp 108-111.
- 10 Rojas Varela, J., "Antennes filtre de polarisation dans la bande HF", thèse d'université, Rennes, 1987.
- 11 Bertel, L., Rojas-Varela, J., Cole, D., Gourvez, P., "Polarization and ground effects on HF receiving antenna patterns", *Annales des Télécommunications*, septembre 1989.
- 12 Le Roux, Y., Nibéron, M., Fleury, R., Menard, J., Jolivet, J.P., "HF channel modelling and simulation", *IEE Conf. Pub.*, n° 325, 1990, pp 72-77.
- 13 Bertel, L., Bertin F., Testud, J., "De la mesure du contenu électronique intégré appliqué à l'observation des ondes de gravité de moyenne échelle", *J. Atmos. Ter. Phys.*, vol. 38, 1978.
- 14 Bertel, L., Le Ray, B., Baltazart, V., Le Saout, J.Y., "Analysis of behaviour of small basis radiogoniometers using interferometry technique", *Conf. on Solar Terr. and Sp. Phys.*, Melbourne 1990.
- 15 Rogier, J.L., Multedo, G., Bertel, L., Baltazart, V., "Ionospheric multipaths separation with a high resolution direction finding algorithm mapped on an experimental system", *IEE Proceedings*, 339, juillet 1991.
- 16 Bertel, L., Rojas Varela, J., "Antennes HF adaptées à la polarisation des signaux reçus : exemples de résultats", *Conf., Procéd., JINA, Nice*, 1986
- 17 Compton, R.T., "The tripole antenna : an adaptive array with full polarization flexibility", *IEEE Transactions on Antennas and Propagation*, vol. AP 29, n° 6, nov. 1981.
- 18 Bertel, L., Rojas -Varela, J., Cole, D., Gourvez, P., "HF antenna behaviour in a receiving system", *IPS Report 87.04*, Sydney, Australie.
- 19 Jeffrey, Z.R., Middleton, P.T., Winkler, C., "Accurate measurements of the total angle of arrival of HF skywaves", *AGARD Symposium*, Munich, mai 1988.
- 20 Gething, P.J., "Radio direction finding and the resolution of multicomponent wave fields", *Stevenage, England*, Peter Peregrinus, 1978.
- 21 McNamara, L.F., "Ionospheric modeling in support of single station location of long range transmitters", *J. Atmos. Terr. Phys.*, vol. 50, 1988, pp 781-795.
- 22 Bertel, L., Gourvez, P., Le Saout, J.Y., Ruelle, N., "HF receiving antenna array behaviour", *IEE Conf. Publ.*, n° 301, vol. 1, 1989, pp 507-511.
- 23 Le Saout, J.Y., Gauthier F., Ruelle, N., Fleury, R., "Utilisation of the Losquet backscatter sounder receiving network using radiogoniometry", *50 th AGARD-EPP Symposium*, Londres, 1-5 june 1992.
- 24 Johnson, D.H., "Multiple source DF signal processing : an experimental system", *IEEE Transactions on Antennas and Propagation*, vol. AP 34, n° 3, 1986.
- 25 Friedlander, G., "A sensitivity analysis of the MUSIC algorithm", *IEEE Transactions ASSP*, vol. 38, oct. 1990, pp 1740-1751.
- 26 Multedo, G., "L'évolution des techniques de goniométrie pour la surveillance radio dans les bandes HF et VUHF : de l'interférométrie aux techniques de goniométrie à haute résolution", *50 th AGARD-EPP Symposium*, Londres, 1-5 june 1992.
- 27 Schmidt, R., "Multiple emitter location and signal parameter estimation", *IEE Transactions on Antennas and Propagation*, vol. AP 34, n° 3, 1986.
- 28 Schmidt, R., "A signal subspace approach to multiple emitter location and spectral estimation", *PHD dissertation*, Stanford University, nov. 1981.
- 29 Ferrara, E.R., Parks, T.M., "Direction finding with an array of antennas having diverse polarizations", *IEEE Transactions on Antennas and Propagation*, vol. AP 31, n° 2, 1983.
- 30 Bienvenu, G., Kopp, L., "Optimality of HR array processing using the eigensystem approach", *IEEE Transactions ASSP* 31, 1983, pp 1235-1248.
- 31 Chen, J., Bennett, J.A., Dyson, P.L., "Automatic fitting of QP segments to ionospheric profiles with applications to ground range estimation for SSL", *JATP*, 52 (4), 1990, pp 227-288.
- 32 McNamara, L.F., "The ionosphere : Communication, Surveillance, and Direction Finding", *Krieger Publishing Company*, 1991.

## DISCUSSION

### C. GOUTELARD

L'utilisation des méthodes spectrales, ne peut être faite, me semble t'il, que sur des émetteurs coopérants ou particuliers. Avez-vous une méthode particulière pour les appliquer à des émetteurs non coopérants?

*It would seem to me that spectral methods can only be used on cooperative or specific emitters. Do you have a special method for applying them to non-cooperative emitters?*

### AUTHOR'S REPLY

Il est exact que les méthodes spectrales nécessitent de travailler en bande étroite et donc sur des émetteurs utilisant une porteuse, qu'ils soient coopératifs ou non. Il en résulte que cette technique sera de préférence utilisée pour des études du milieu de propagation ou d'évaluation d'autres techniques.

*It is true that spectral methods will only work in a narrow band, and therefore on emitters, which use a carrier, whether they are cooperative or not. Therefore, this technique will preferably be used for studies of the propagation environment or for the evaluation of other techniques.*



# A CURRENT ASSESSMENT OF SINGLE SITE LOCATING TECHNOLOGY

Robert B. Rose  
 Ionospheric Branch, Code 542  
 Ocean and Atmospheric Sciences Division  
 Naval Command Control Ocean Surveillance Center, RDT&E Division  
 San Diego, CA, 92152-5000, USA

## ABSTRACT

This paper describes the observed performance of a HF Single Site Locator (SSL) which employs current interferometer technology. 142 locations were obtained from non-cooperative targets over a five day period in April 1989. Locations were determined from fast, high resolution azimuth and elevation angle of arrival measurements and a knowledge of the ionospheric reflecting medium. Exact emitter locations were determined post facto with ground truth data from the field units. Five modes of transmission were encountered and the signals were successfully located. Miss distance accuracies varied between 12.5 kilometers to 40.5 kilometers over ranges that varied from 109 kilometers to 526 kilometers. Performance varied as a function of modulation type with SSB voice being the most difficult to prosecute and packet and burst signals were the easiest to locate. Also performance varied as a function of range and the relationship of the operating frequency to the maximum usable frequency between the SSL and the target.

## I. INTRODUCTION

HF radiolocation will remain important to the surveillance community through the next several decades. Use of the HF spectrum continues to expand, especially in third world countries. The number of emitters will increase and the use of more exotic modes of transmission is also expanding. From the surveillance point of view, the HF signal location scenario for the next decade is one in which high interest regions will change as new crises arise. The areas where signals of interest (SOI) will originate will be smaller so signal search can be more focussed. A key element in future direction finding is expressed by the term "rapid deployment." Smaller, tactical systems that are easily transportable will replace the large shore-based direction finding networks.

Single Site Location (SSL) technology was first conceived in the late 1960s. However, its initial development was overshadowed by the large shore-based and smaller shipboard netted systems. In those systems, signals were located by triangulation using multiple lines of bearing. But recent events have revealed that the present netted-sensor concept lacks mobility and flexibility, both of which are noteworthy attributes of SSLs. In the 1970s, early testing burdened SSL technology with a "10% of Range" accuracy specification. And in the early 1980s, systems developed and tested by the research community were lacking in performance. Furthermore, a lack of awareness of the real reasons for this

breakdown in SSL performance compromised their further advancement. Tests had indicated that signal sampling rates had to be higher and propagation effects had to be accounted for. Attempts to develop the next generation system failed to capitalize on history, and since then SSL as been, in general, ignored as a practical radiolocation technology by the majority of the HF surveillance community.

Nevertheless, recent advancement of SSL capabilities has been achieved through cooperative programs and in-house research. This paper describes the results of one such exercise conducted by personnel from the Naval Command Control and Ocean Surveillance Center (NCCOSC), RDT&E Division (formerly, Naval Ocean Systems Center) in San Diego, California in 1988.

Transionospheric refractive effects were measured at 30 MHz by comparing the direct elevation angle of arrival of signals from an orbiting HF beacon with the accurately known altitude/location of the satellite. The angle of arrival measurements were made with the Single Site Locating Testbed at Southwest Research Institute (SwRI) in San Antonio, Texas. The transionospheric work had required upgrades and calibration of the 7-channel interferometer which had been originally designed as an SSL. It was desired to demonstrate that these modifications had improved the SSLs capabilities against terrestrial HF signals as well.

In April 1989, the exercise described here was conducted with the SwRI SSL. A completely independent communications exercise was being conducted in Texas by another facility which provided numerous, different types of signals. Many signals of opportunity offered a good chance to evaluate the performance of a modern single site locator in a non-cooperative signal environment. Modulation types were a mix of older, more familiar modes and newer modern digital schemes. This also provided the first opportunity to make SSL measurements of digital burst and packets signals. The data were considered non-cooperative in that the locations of the SOIs were not known in advance of the measurement period. Also their transmission schedules were not known.

## II. DESCRIPTION OF THE SINGLE SITE LOCATING TESTBED

Radiolocation with a single station location system includes an azimuth measurement on the target's signal and an estimate of the target's great circle range based on ionospheric parameters measured at the SSL site. The system combines azimuth and range to the target to produce the targets location. This is also expressed in

the corresponding geographic coordinates as illustrated in figure 1.

Figure 2 illustrates the essential concept of SSL system operation. An HF radio wave transmitted from a target transmitter reflects from one or more reflecting regions in the ionosphere. The radiowave propagated by each mode arrives at the SSL site with a given bearing and angle of elevation. A radio direction finder measures the azimuth and elevation of the arriving signal.

The SSL system includes an ionospheric sounder which measures the virtual height of the ionospheric reflecting region. The reflection height measured at the SSLs location is assumed to be the same as at the reflection midpoint. As shown in figure 2, the triangle from the SSL site through the ionospheric midpoint and down to the target transmitter provides a location estimate. This calculation generates a range estimate (in kilometers) and combines it with the observed azimuth bearing. The SSL system output is thus an estimate of the location of the target producing the signal expressed in geographical coordinates. This is in contrast to a conventional direction finder which provides only an azimuthal line of bearing to the target signal.

The range calculation uses the measured elevation and virtual height of the ionosphere, along with the law of cosines, to determine the angle at the center of the earth from the midpoint (half range) to the SSL system. This angle, when multiplied by twice the earth radius, gives the great circle range from the SSL site to the target.

Figure 3 shows the fundamental system components of the SwRI SSL system used for the tests described in this report. The computer interfaces both to the phase-measuring interferometric direction finder and to the ionospheric sounder. The antenna system has two parts, a low band array (3-10MHz) which is 150 meters on a leg and a high band array (10-30 MHz) which is 50 meters on a leg. Both arrays are in an "L" configuration using seven crossed loop elements. The azimuth and elevation angles of an incoming HF signal were calculated using a phase linear concept. SwRI refers to this concept as a coincidence direction finding interferometer. In this type of system, the phase linearity is tested on the resolved long baseline phase for each leg of the SSL antenna array. If the differences between the predicted and measured value on both baselines are less than some preset value, then the data frame becomes a coincidence frame, one in which the array has resolved the angle of arrival (AOA) of the signal.

In single site locating, operator functions were to search and identify a signal of interest, optimize receiver tuning, and start and terminate the direction finding process. Automatic functions of the system include computing the target azimuth and elevation angles of arrival, measuring the ionospheric height, generating an elevation angle/range transmission curve and computing the best

point estimate (BPE) of the target location. Figure 4 shows the operator display on the present SSL Testbed. Once the operator has selected a SOI, he presses the "execute" button initiating the scanning process on the array. The azimuth and elevation are displayed for every frame that passes the phase linearity test. When the operator decides there are sufficient data to calculate a fix, the computer is switched to the "location" mode. Then the elevation cursor is placed on the range transmission curve to achieve a "range" solution. The azimuth angle of arrival calculation is automatic and requires no operator intervention. After completing the calculation of the range and azimuth angle, the computer generates the target's geographical coordinates. There is no way for the system to distinguish between a one- and a two-hop signal. Therefore the system generates both solutions, leaving the decision up to the operator.

In the past, there have been three issues which have a significant impact on the accuracy of an SSL. These are (1) the assumptions made in the BPE calculation; (2) the assumption that the signal propagates along the great circle path at short ranges; and (3) the interval at which the ionospheric data needs to be updated.

The assumptions made in the BPE calculation lead to three weaknesses. The assumptions are: (1) that the operator will place the elevation cursor on the correct elevation angle and will do so "properly"; (2) that the transmission curve calculation, generated from the vertical ionograms, is correct and current; and (3) that the ionospheric height at the path midpoint is approximately the same one measured locally.

Near vertical incidence propagation, refers to skywave propagation between 100 and 500 kilometers. It has long been thought that near vertical incidence signals inside 100 kilometers will not produce reliable AOA information. A primary assumption in SSL is that the ionospheric reflection point is along the great circle path at the midpoint. Experimental evidence (Paul, 1985) indicates that the skywave reflection point wanders from several kilometers to tens of kilometers from the great circle path. For ranges greater than 500 kilometers, this represents a negligible error whereas inside 500 kilometers errors are significant and inside 100 kilometers they are intolerable. Also for ranges of less than 100 kilometers where the elevation angles are greater than 82 degrees, the arctangent function for the angle of arrival equations becomes very unstable. While it is recognized that tilt correction techniques exist, they have yet to be fully demonstrated in a realistic operating environment.

The ionospheric profile has to be updated at a rate which will maintain acceptable SSL accuracy because the maximum correlation time between ionospheric measurements is 5 minutes (Rose, 1988). One objective of this experiment was to see whether increasing the number of ionospheric measurements over a given period of time had any impact on the accuracy of the fix.

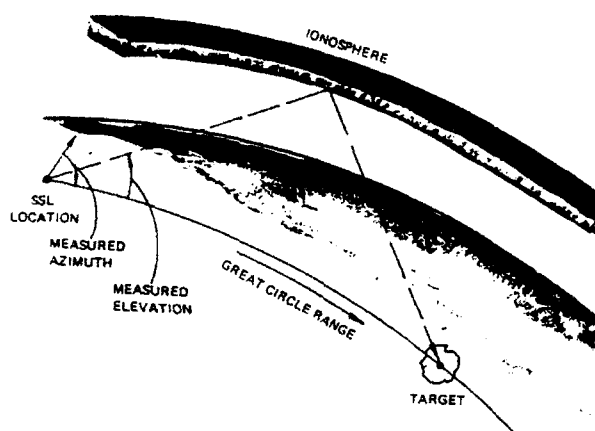


Figure 1. Single Site Locating Concept

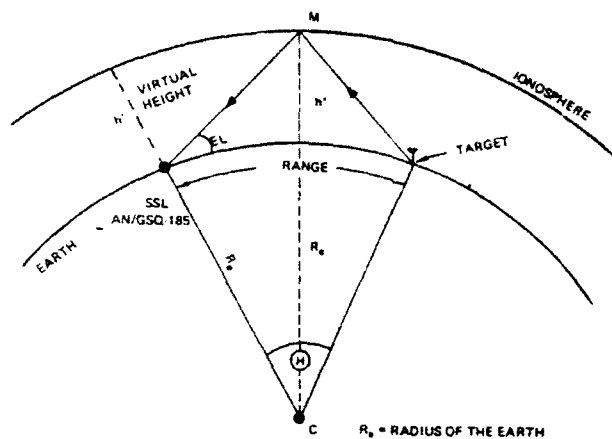


Figure 2. Range Estimate Geometry

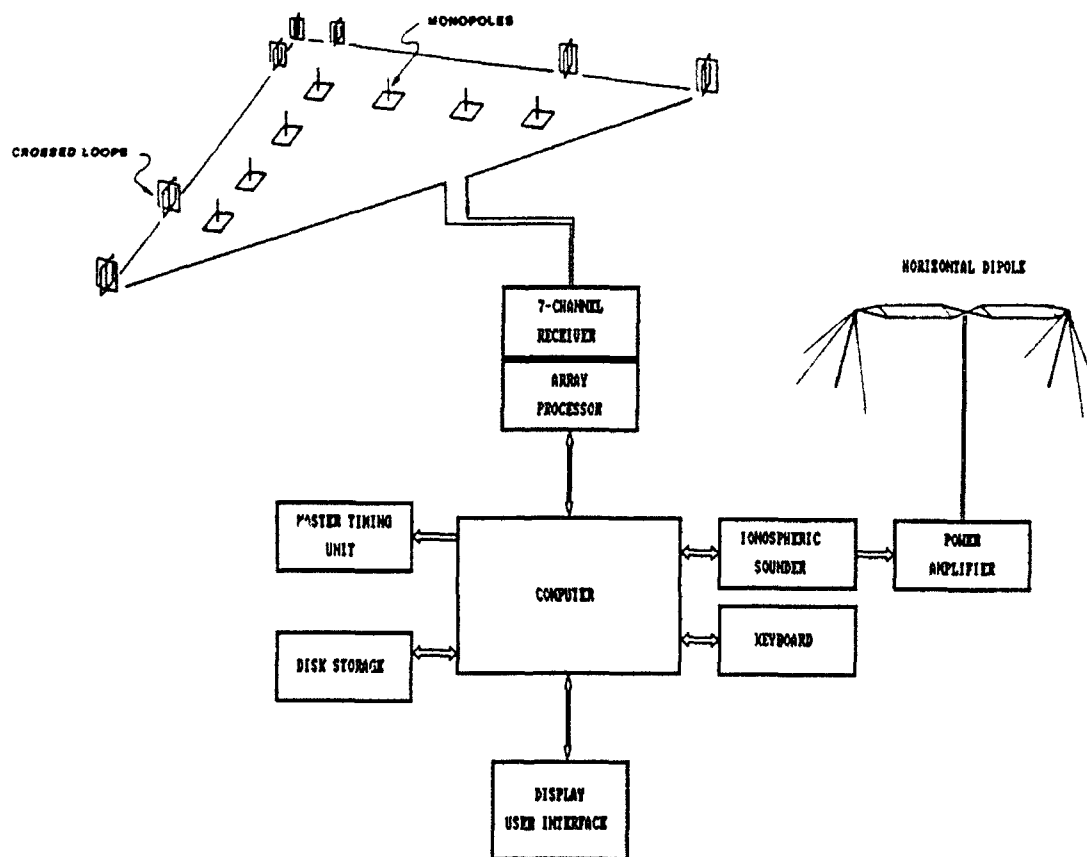


Figure 3. 7-Channel Single Site Locator Instrumentation

### III. DESCRIPTION OF THE FIELD TEST DATA COLLECTION

Data collection for this test occurred between 20 and 25 April 1989. During this period, solar conditions were relatively undisturbed. The 10.7 cm flux varied between 173 and 198. The magnetic A-index was low, ranging from 8 to 14. At midday, the F-region critical frequencies were exceeding 15 MHz which means that the 3000 km maximum usable frequency exceeded 50 MHz. Generally propagation conditions were excellent.

The frequencies used were known before hand. The operating schedule and the mode of transmission were unknown. Thus, scanning the allocated frequencies was required to locate the signals which were part of the experiment. Ground truth location data could be obtained after the fact. While this approach may appear rather "hit or miss," only a day of monitoring was required to establish the frequencies used and the general locations of all the participants in the experiment.

The five day collection period produced locations for 142 signals. After finding a SOI, the system was activated. Throughout the test period, the ionosphere was measured and data processed automatically and stored in the computer data files. A key issue was the speed with which the SSL could acquire a signal, gather data and generate a fix. After learning the mechanics of the testbed, the entire process could be accomplished in a matter of seconds. At no time were there more signals than the testbed could prosecute, and all signals prosecuted produced a fix. All of the data collected were reported. The SSL testbed was never stressed by saturation. In fact a major concern was the relative lack of signals to prosecute. An entry for the signal call sign, event number and signal characteristics for each signal prosecuted were manually entered in the operator log. When the computer provided the BPE for the event, the hardcopy output was also taped in the log as was all the related data for each event. These data were then used to reconstruct skywave propagation conditions that could influence SSL accuracy.

After the field test, "ground truth" data were obtained on the locations of the mobile and fixed emitters. With these data, and the information in the operator's log, casecards were constructed for each of the 142 fixes. The casecard software, written during the 1986 experiments, allows the user to determine the accuracy of the SSL fix product. For this exercise, the casecard data were entered into Lotus 123 spreadsheets for analysis. The next section presents results from these analyses.

### IV. PRESENTATION OF THE DATA

The experiment produced 142 location fixes on signals that varied in range from 95 kilometers to 528 kilometers. Before this exercise there was very little quantitative data on SSL performance on short path (under 500 Km) skywave signals. The experiment encountered five different modulation types:

- (1) Four Second Digital Burst - 40 fixes
- (2) Digital Packet (encrypted) - 29 fixes
- (3) Upper Sideband (USB) Voice - 24 fixes
- (4) Amplitude Modulation Voice - 28 fixes
- (5) Morse (18 groups/minute) - 21 fixes

In addition to SSL accuracy, other issues to be studied during this exercise, were operator experience, display interpretation and the user interface. The operator had a broad experience in SSL development, direction finder operation and HF propagation and modeling. The accuracy of the present SSL concept relies heavily on operator judgement and his understanding of what is being displayed. It was desired to gain enough experience with these issues in mind to develop expert systems to automate some of these processes in the future.

An early observation in this experiment was that each signal modulation type produces very different data clustering on the operator display. This in turn affects the difficulty in placing the range cursor and the accuracy of the location estimate. Figure 5 shows examples of four of the modulations encountered and the type of presentation the operator had to contend with.

The encrypted digital packet signals were very easy to find, even in a high interference environment. These signals were also easy to fix and produced a very high phase linear ratio. However, they all came from a single location 95 kilometers away and the short range caused severe azimuth spreading. This badly corrupted the results from this one site.

By far the most difficult conventional narrowband signal to prosecute with the present system is single sideband, suppressed carrier (SSBSC) voice. Because the received power varies as a function of the speech envelope, the signal presents a variable signal to noise, to which the system must constantly be adjusting. Very low (below 20%) phase linear ratios are quite common with this mode of transmission. For this reason, SSB signals normally require 30-60 seconds of collection time to acquire a sufficient sample to process. It is very frustrating for a SSL operator to listen to a signal that is very loud and very easy copy to the ear and yet see the phase linear sample count not increasing. This could be rectified with more sophisticated processing in the next generation system. The example in figure 5 shows an unusually good example of upper sideband (USB) voice as it appeared to the SSL operator. Normally the displays were much sparser.

Amplitude Modulation (AM) voice signals provide a continuous carrier which provides in a stable signal to noise ratio and therefore is a relatively easy signal for the SSL to prosecute. Normally, 10 seconds will provide a phase linear sample sufficient to locate the signal. Figure 5 illustrates how the AM signal appears on the SSL display. In this example it took 10 seconds to acquire 487 samples.

Figure 5 also shows how on-off keyed

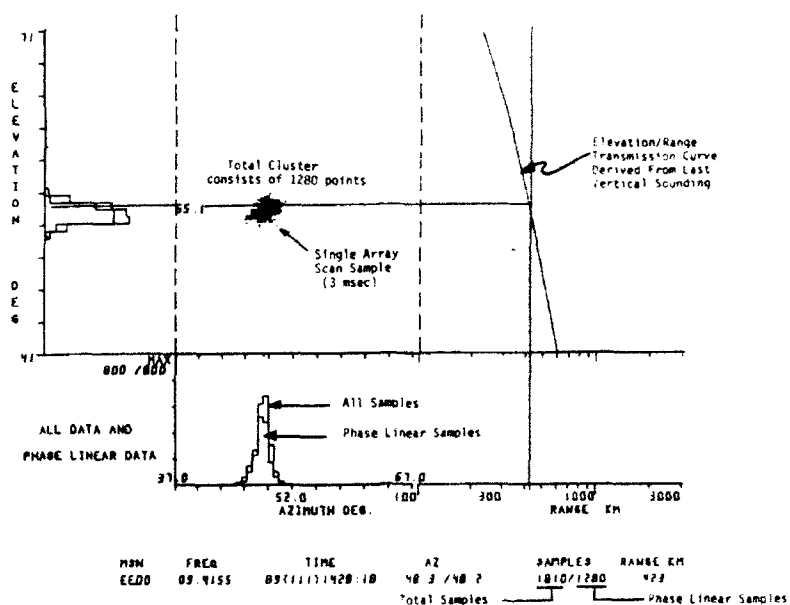


Figure 4. SSL Azimuth/Elevation Operator Display for an USB Voice Signal

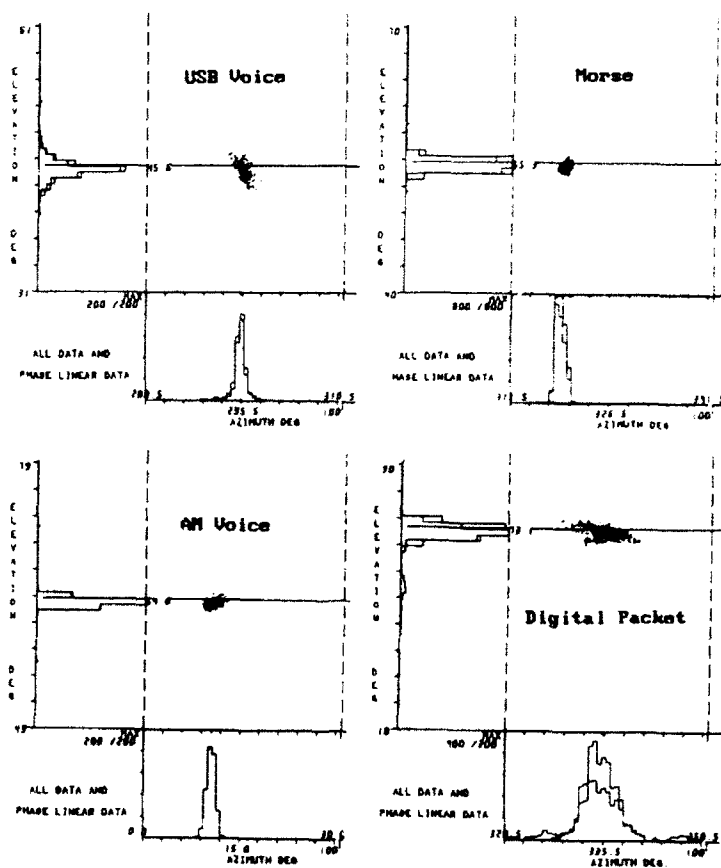


Figure 5. Scope Presentations for Different Modulation Modes

Morse signals appeared to the SSL test bed. In the example shown it took 90 seconds to acquire approximately 2000 samples. At reasonable on-off keying speeds (i.e. 15-25 code groups per minute), this mode of transmission presents a very stable signal to noise making it one of the easiest signals to locate. Experience has shown that any part of a morse transmission produces a usable location product. In most cases the accuracy of the 3 second fix was the same as the 90 second fix. It was possible to get good fixes on single dots or dashes.

Table 1 summarizes the SSL performance by modulation type. One commonly used measure of SSL performance is the miss distance, the distance between the BPE and the true location of the target. This can be expressed in linear distance (kilometers in this case) or as a percentage of the true range between the SSL and the target. In this latter case, the performance measure normalizes out the differences in baseline ranges. Review of all of the fix data indicates that 26 (18%) of the fixes had miss distances of less than 10 kilometers, 78 (52%) had miss distances of less than 20 kilometers and 106 (75%) had miss distances of less than 30 kilometers. This is quite good considering the short ranges of the data used for this study.

Figure 6 shows how location azimuthal accuracy degrades as range decreases. A signal from 423 kilometers presents an azimuthal dispersion of approximately 6 degrees whereas a signal from 160 kilometers is spread over 30 degrees. The longer range produced a better location fix, one with a lower variance. On the other hand, the very short range produced a location fix with a very high variance which can be interpreted as very low confidence in the result.

## V. DISCUSSION

The exercise described in this report provided new insight on two digital signals, the short burst and the packet, that had never been prosecuted before with this SSL and short baseline ranges of between approximately 100 and 500 kilometers. The measurements resembled a tactical situation against uncooperative signals whose location was unknown at the time. While the upgraded seven channel interferometer samples signals and makes an AOA measurement very rapidly, the user interface was not modernized and presented a less than optimum situation. Although the operator for this test had extensive experience in the areas of HF direction finding, SSL development and HF

propagation, interpretation of the display was still often complicated. Therefore the performance results, while considered good, can be improved.

The four second digital burst communication mode provides the interferometer with an optimum signal. This mode is extremely easy to prosecute and, at ranges of greater than 300 kilometers, it is easily located. As the range decreases below 300 kilometers, that accuracy of SSL fixes degrades. Indeed, at these the shorter ranges (<300 km), such signals are actually harder to locate accurately than signals with a higher duty cycle. Because of their short duration, the AOA from a burst signal is vulnerable to any shifts in the reflection point from great circle. The result presents the user with an easily heard and usually loud signal, as well as optimum azimuth and elevation displays with very small standard deviations and a definite, but wrong location solution.

During the experiment, the time between ionospheric soundings was varied. At the times of the day when the sun was low in the sky, updates were needed every 10 minutes to keep fix accuracies between 5% and 10% of range. During midday and at night, sounding every fifteen minutes was sufficient. When the ionospheric data were updated hourly, the performance degraded to 10% to 15% of range. Historically, hourly updates have been the norm.

## VI. CONCLUSIONS

Single site radiolocation at HF is a good technology for signal sources located between 100 to 1500 kilometers. Future conflicts will require a quickly deployable intelligence gathering system in which radiolocation is an important function. SSL can do this. Over the last several years, the new improvements in various components in this technology have been demonstrated by one means or another. It is now time to design a well thought out system that takes advantage of the latest advances in computer systems, intelligent user interfaces using expert systems and modern ionospheric sensing.

The data reviewed indicate that the performance of the SSL is influenced by the transmission mode. The mean miss distances varied from 12 kilometers to 40 kilometers which represents a variation between 6.0% and 11.0 % of range. It is estimated that a next generation SSL can achieve 5.0% of range accuracies if the user interface and the vertical sounder are modernized and

Table 1. Summary of SSL Accuracy Performance

Modulation	Mean Miss Distance (km)	Mean Miss Distance (% of Range)	Mean Azimuth Error (Deg)
Burst	22.3	6.0	1.1
Packet	12.5	8.3	-0.3
USB Voice	40.5	11.0	1.1
AM Voice	24.0	7.7	-1.3
Morse	23.4	7.4	-1.8

made smarter. It must incorporate more collateral information about the intended targets. Finally, the next generation system needs some sort of propagation mode identification.

The experimental data presented in this report provides a glimpse of how good single site locating can be today and also, provides direction for improvement. An earlier HF radiolocation experiment established that the baseline accuracy of the SWRI seven channel interferometer testbed is 1.0% of range without the errors induced by the ionosphere. This is likely to be close to the physical limit of a SSL when it is prosecuting terrestrial targets since the earlier measurements were made just before sunrise when the ionospheric electron density is at its minimum.

The requirement of high quality real time vertical sounding data for the range solution is critical to the SSL process. This experiment confirmed earlier work that the performance of the system is heavily dependent upon (1) the experience of the operator and (2) the use and interpretation of the user interface. SSL performance has been in the past and remains today very "user intensive." The next generation SSL should have an "intelligent" user interface to assist in the data interpretation and decision making.

The ionospheric sounding process needs the following improvements: (1) ionospheric sensing should be sped up, updating the ionogram every two minutes. Experimental evidence indicated that the ionosphere is more variable than current models show; (2) the transmitted signal use spread spectrum techniques to reduce interference; (3) the ionogram interpretation and range/elevation angle transmission curve needs to be modernized through the use of an expert interpreter system; (4) the sounder operation and interpretation must be transparent to the user.

#### ACKNOWLEDGEMENT

The author acknowledges the contributions of Dr. D.R. Lambert in the review and editing of this paper.

#### REFERENCES

- Rose, K.B., "High-Resolution HF Time of Arrival measurements (1981-1985)," Radio Science, Vol 23, Number 3, May-June 1988
- Paul, A.K., "F-Region Tilts and Ionogram Analysis," Radio Science, Vol.20, Number 4, July-August 1985

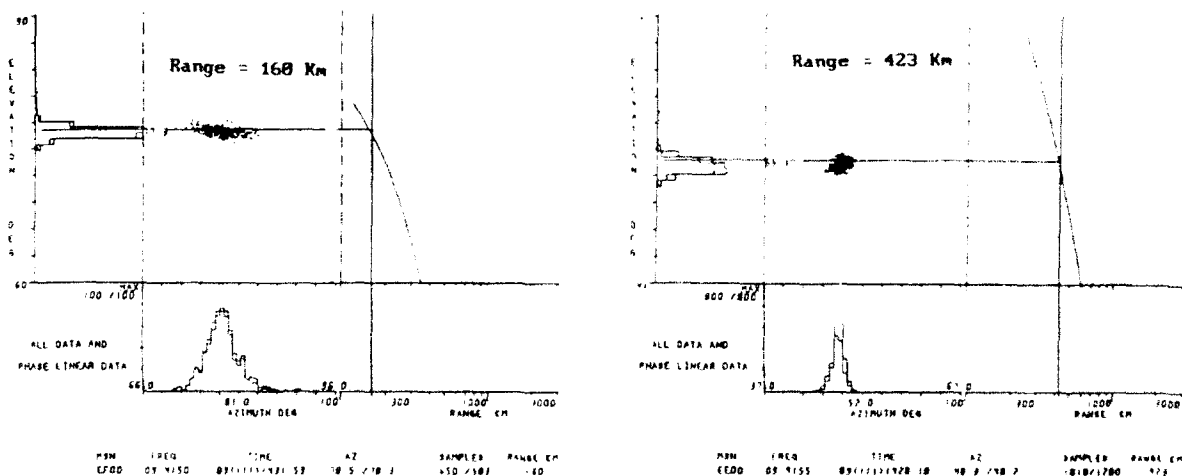


Figure 6. Comparison Between a Very Short and Medium Range SSL  
Fix. Modulation Mode is a Four Second Digital Burst.

## DISCUSSION

### H. SOICHER

Since ionospheric conditions are updated by a vertical sounder, what assumptions do you make about horizontal variations of the ionosphere?

### AUTHOR'S REPLY

It is assumed, based on work by Charles Rush, that the correlation distance of an ionosonde measurement is roughly 500 km which practically limits it to tactical targets with ranges of less than 1000 km. For the test conducted, the target ranges were between 100 and 500 km, or ionospheric reflection points between 50 and 250 km. At these ranges, I was very comfortable using a local measurement.

### G. HAGN

Your data (e.g., Figure 6) illustrate the problem of obtaining good range estimates with the single site location (SSL) method with the interferometer technique on short paths. You mentioned that this method is not practical for ranges less than 100 km without employing a tilt correction. Even with a tilt correction, there may be limitations at short ranges. The MUSIC algorithm, when used with suitable antennas (e.g., the CART antenna described in the Proceedings of the 1990 Tactical Communications Conference, Fort Wayne, IN) can be used to estimate the angle of arrival (AOA in azimuth and elevation). Of course, it is still necessary to have good data on the ionosphere and a good ray tracing model to invert these AOA data to estimates of emitter location. It would be interesting to test the MUSIC approach on short paths (in the steep part of the Ross curve) versus the interferometer method with tilt correction.

### AUTHOR'S REPLY

While what you say might be true, I am skeptical about using the MUSIC algorithm in this manner. I have observed the application of MUSIC in many applications over the last decade, and those I have observed have not resulted in a practical, "real world" solution. I should say, I have yet to be convinced that dynamic tilt correction inside 100 km is practical. I have read most of McNamara's and Georges' papers on the subject. Translating these ideas into a practical, useable system is still developmental and not a proven fact. I think the work done so far shows we have a "real" SSL capability at tactical ranges of 1000-100 km. To attempt to do accurate skywave DF inside 100 km (over land) may not merit the added complication.

### S. TOWNES

How well did you know the antenna calibration and in an operational system how effective is the calibration?

### AUTHOR'S REPLY

The SWRI interferometer antenna system was very tightly calibrated using a helicopter point source. We know the patterns very well. In developing the calibration, experience was gained in learning how the antenna patterns degrade in field operation. If the field interferometer is set up to specification and the heights of each element are accurately known, the antenna degradation ought not be more than 5%.

### LEROUX

Quel est l'intérêt du réseau de monopoles? Comment traitez-vous les multirebonds?  
What is the benefit of a monopole array? How do you deal with multiple hops?

### AUTHOR'S REPLY

The monopole provides uniformity in the antenna pattern. As with any SSL, one must assume single-hop propagation.

### J. BELROSE

Would you comment on how to make a covert vertical incidence sounder?

### AUTHOR'S REPLY

Actually, there are two methods. First, it is easy to modify the frequency synthesizer to follow a pseudo-random hopping sequence. Instead of a linear sweep, the sounder hops until the ionogram is constructed. We have tried this and it works. Second, we are testing 40 kHz spread-spectrum signals which are undetectable. There is no reason this couldn't be swept in frequency to measure ionospheric height.



## A NEW PASSIVE SSL TECHNIQUE FOR HF RADIOLOCATION

R. L. Johnson, Q. R. Black and A. G. Sonstebj  
Electromagnetics Division  
Southwest Research Institute  
6220 Culebra Road  
San Antonio, TX 78238, USA

### SUMMARY

This paper considers the problem of passive radiolocation for the case of HF multipath propagation. A new technique is developed for the estimation of interpath time delay applying an eigen based super-resolution spectral estimation method. The technique samples the wavefield received by two spatially separated antennas to compute intersensor delay time for each constituent signal through a normalized eigen cepstral analysis. The intersensor delays and interpath delay are used in concert to estimate the location of the transmitter. The method is applied to experimental data in a preliminary proof-of-concept analysis.

### 1. INTRODUCTION

In this paper we consider the problem of locating a radio transmitter through measurements made from the wavefield received at a single site. The technique used to solve this problem is generally referred to as single site radiolocation (SSL) [1-4]. Although this is an area of active research in several disciplines, we consider mainly HF (3 - 30 Mhz) communications and ionospheric propagation. The conventional approach in the SSL process is to estimate the angle-of-arrival (AOA) of the radio signal at one receiving site, the AOA being characterized by azimuth and elevation. Next one estimates the height of the ionospheric reflecting layer, typically accomplished by a vertical or oblique incidence sounding. Ground range to the transmitter is computed using the estimated elevation AOA and ionospheric height. Transmitter location is determined from the estimated azimuth AOA and ground range.

In contrast to the conventional SSL technique described above, multipath propagation through the ionosphere can be exploited to accomplish passive radiolocation without requiring height data from a sounder. The essential approach in passive location is to sample the received signal and to estimate the interpath delay time. There are three published works which are of special relevance to the approach we develop in this paper. The first work is that of Carter, et. al. [5] and their description of the smoothed coherence transform (SCOT). The SCOT was developed

to estimate time delays between weak broadband correlated noise received at two sensors. The technique first computes the complex coherence function between the two sensor outputs and then applies the fast Fourier transform (FFT) to compute the inverse discrete Fourier transform of the estimated coherence. The result is a sharp peak indicating the time delay between common processes contained in the two sensor outputs. A second work which particularly relates to our approach is the use of interpath delay time to compute ground range as described by Gething [1]. The technique uses the differential group delay of two ionospheric propagation paths and the corresponding elevation AOA's to compute range. The approach is an application of Breit and Tuve's theorem [6] for a thin layer ionospheric model. The third work is the location performance analysis done by Rendas and Moura [7]. They consider the general problem of passive location using intersensor and interpath delays and derive the Cramer-Rao bound for the errors in the location estimate.

In this work we propose a new approach to the passive SSL problem which is a merger of the complex cepstrum and superresolution techniques. The conventional procedure used to compute the complex cepstrum is to perform an FFT on a time series which contains one or more echos. A complex logarithm is taken of the result to "whiten" the data, and an inverse FFT is computed to estimate the echo delay times. Our approach differs from the conventional technique in several respects. Similar to the approach described by Carter, et.al. [5], we consider *two* spatially separated time series and compute a normalized cross power spectrum (or complex coherence function) to remove modulation effects. The cross power spectrum is parametric in delay time between sensors for each signal and also the differential path delay time from the transmitter. To determine the intersensor and interpath delay times, we do not take the logarithm but rather compute a spectral cross correlation matrix and apply the MUSIC [8] spectral estimation technique. We describe this process as the normalized MUSIC cepstrum. The location of the transmitter is estimated by applying Breit and Tuve's theorem to the intersensor and interpath delay data.

## 2. NORMALIZED MUSIC CEPSTRUM

Consider a transmitted signal propagating via a groundwave and an ionospherically reflected skywave which is received by two antennas. (Although the multipath geometry is completely arbitrary, we have chosen a groundwave/skywave combination for illustration.) We characterize the transmitted signal as  $s(t)$  and denote the differential group delay time for the signal to propagate via the two paths as  $\tau_m$ . Also we characterize the propagation delay between the antennas in the receiving array as  $\tau_g$  for the groundwave and  $\tau_s$  for the skywave. The signal received at each antenna may be expressed as

$$x_1(t) = A_g s(t) + A_s s(t - \tau_m) \quad (1a)$$

$$x_2(t) = A_g s(t - \tau_g) + A_s s(t - \tau_s - \tau_m) \quad (1b)$$

where  $A_g$  denotes the group path attenuation for the groundwave, and  $A_s$  the attenuation for the skywave path. If  $F(\omega)$  denotes the Fourier transform of  $s(t)$ , then the Fourier transform of  $x_1(t)$  and  $x_2(t)$  may be represented as

$$F_1(\omega) = [A_g + A_s e^{-j\omega\tau_m}] \cdot F(\omega) \quad (2a)$$

$$F_2(\omega) = [A_g e^{-j\omega\tau_g} + A_s e^{-j\omega(\tau_s + \tau_m)}] \cdot F(\omega) \quad (2b)$$

The normalized cross power spectrum between receiving antennas is given by

$$S_{12}(\omega) = \frac{F_1(\omega) F_2^*(\omega)}{\sqrt{|F_1(\omega)|^2 \cdot |F_2(\omega)|^2}} \quad (3)$$

The normalization is applied to reduce the effects of signal modulation. To express the normalized cross power spectrum in polar form, we simplify the notation by defining  $A_1$  and  $A_2$  as,

$$A_1 = A_g + A_s e^{-j\omega\tau_m} \quad (4a)$$

$$A_2 = A_g e^{-j\omega\tau_g} + A_s e^{-j\omega(\tau_s + \tau_m)} \quad (4b)$$

The normalized cross power spectrum may alternately be expressed as

$$S_{12}(\omega) = \frac{|A_1 A_2^*| \cdot \arg(A_1 A_2^*) \cdot |F(\omega)|^2}{\sqrt{[A_1 A_1^* |F(\omega)|^2] \cdot [A_2 A_2^* |F(\omega)|^2]}} \quad (5a)$$

$$S_{12}(\omega) = \text{Arg}(A_1 A_2^*) = e^{j\phi(\omega)} \quad (5b)$$

As a result of the normalization process, the cross power spectrum has unity magnitude and is characterized by a complex argument which has a periodicity determined by the intersensor delays and the interpath delay. The amplitude of the periodicity is a function of the relative amplitudes of the two propagation paths. To estimate the intersensor and interpath delay times, one must accomplish a spectral decomposition of the cross power spectrum, and to do this, we apply the MUSIC superresolution technique. The MUSIC method is employed because the signals of interest are generally contained in a narrow bandwidth, and as a consequence, the number of periods which we may observe is usually restricted. Also the estimate of cross power spectrum is generally contaminated by noise and may be expressed as

$$S_{12}(\omega) = e^{j\phi(\omega)} + N(\omega) \quad (6)$$

where  $N(\omega)$  may be a complex expression involving (signal  $\times$  noise) and (noise  $\times$  noise) terms. In this development, we consider the case of additive white Gaussian noise which is uncorrelated with the signal components.

Consider the situation in which the time series at each antenna is sampled and an FFT is computed for each antenna output. For each frequency bin which contains signal power, we compute a normalized cross power spectrum defined by Equation (6). If the signal energy were spread over  $M$  frequency bins, then the result would be,

$$S_{12}(\omega) = \begin{bmatrix} e^{j\phi(\omega_1)} \\ e^{j\phi(\omega_2)} \\ \vdots \\ e^{j\phi(\omega_M)} \end{bmatrix} + \begin{bmatrix} N(\omega_1) \\ N(\omega_2) \\ \vdots \\ N(\omega_M) \end{bmatrix} \quad (7)$$

Using the conventional outer product definition, we form the spectral cross correlation matrix as  $C(\omega) = E\{S_{12}(\omega) S_{12}^H(\omega)\}$ , where  $E\{\}$  is the statistical expectation and  $H$  denotes conjugate transpose. The elements of the  $C$  matrix are given as

$$c_{ij}(\omega) = E\{[e^{j\phi(\omega_i)} + N(\omega_i)][e^{-j\phi(\omega_j)} + N^*(\omega_j)]\} \quad (8)$$

$$= e^{j[\phi(\omega_i) - \phi(\omega_j)]} + \sigma^2 \delta(\omega_i - \omega_j)$$

In order to estimate the delay times, we perform a spectral decomposition of the  $C$  matrix using the MUSIC superresolution technique.

$$\phi(\omega) = \tan^{-1} \frac{A_g^2 \sin \omega\tau_g + A_s^2 \sin \omega\tau_s + 2A_g A_s \cos \frac{\omega}{2} (2\tau_m + \tau_s - \tau_g) \sin \frac{\omega}{2} (\tau_g + \tau_s)}{A_g^2 \cos \omega\tau_g + A_s^2 \cos \omega\tau_s + 2A_g A_s \cos \frac{\omega}{2} (2\tau_m + \tau_s - \tau_g) \cos \frac{\omega}{2} (\tau_g + \tau_s)} \quad (5c)$$

To apply the MUSIC algorithm, we form the spectral cross correlation matrix  $C$  and perform an eigen decomposition. The eigenvalues are ordered in monotonically decreasing order as  $\lambda_M \geq \lambda_{M-1} \geq \dots \geq \lambda_1$ . The corresponding eigenvector matrix is  $C_E = [E_M \ E_{M-1} \ \dots \ E_1]$ . For the case of two coherent paths,  $C_E$  will be rank one, and will have a nullspace of dimension  $M-1$ . In general, the rank of  $C_E$  will depend upon the number of coherent path pairs. For one pair of interfering paths, the nullspace matrix is given as  $C_N = [E_{M-1} \ E_{M-2} \ \dots \ E_1]$ , and we compute the MUSIC spectrum as,

$$P = (B^H C_N C_N^H B)^{-1} \quad (9)$$

where  $B$  is the "steering vector" defined as

$$B = \begin{bmatrix} \exp j\phi(\omega_1) \\ \exp j\phi(\omega_2) \\ \vdots \\ \exp j\phi(\omega_M) \end{bmatrix} \quad (10)$$

and the complex argument  $\phi(\omega_i)$  is given by Equation (5c). The essential procedure for computing the delay times is to search for the peaks of Equation (9). If we normalize the amplitudes, the search is through a four dimensional space, namely  $A_g/A_s$ ,  $\tau_g$ ,  $\tau_s$ , and  $\tau_m$ .

### 3. SEA PATH EXPERIMENT

As a preliminary proof-of-concept experiment, an HF transmitter was placed onboard a ship in the Gulf of Mexico, and a receiving site was instrumented on the coastal barrier island near Port Aransas, Texas. The ship was deployed at a distance of 225 Km and azimuth of 89.1 degrees relative to the receiving site. The receiving antenna array geometry is shown in Figure 1. Four crossed loop antennas were located at the positions marked 1 through 4, and the diameter of each loop was 1.524 m (60 in). The transmitter was operated at 8.4Mhz so that one wavelength was 35.7 m. The direction of propagation was along the longest array dimension of 150 m or  $4.2\lambda$  (antennas 1-4).

The four baselines used to compute the interpath delay time and the intersensor delay times were the 1-2 ( $2.9\lambda$ ), 1-4 ( $4.2\lambda$ ), 2-3 ( $1.7\lambda$ ) and 2-4 ( $2.5\lambda$ ). Using an linear sweep FM sounder on the ship, it was determined that the two propagation modes which existed at 8.4 Mhz were a surface wave and an F-layer ionospheric skywave. The signal transmitted for the test was a multitone frequency shift keyed format which filled the 3Khz passband of the receivers. The sampling rate of the analog signal was 28.888Khz so that an array snapshot was acquired every 34.616  $\mu$ sec. Each normalized cross power spectrum was computed using a 512 point FFT from 1024 time samples. An averaged spectral cross correlation matrix

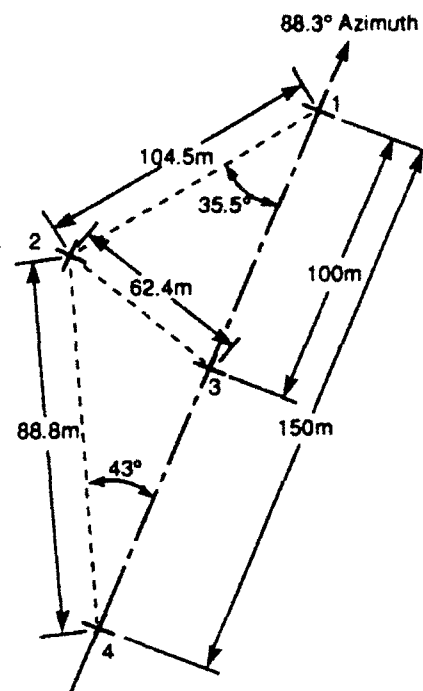


Figure 1. Ground Deployment of Receiving Antenna Array for Sea Path Experiment

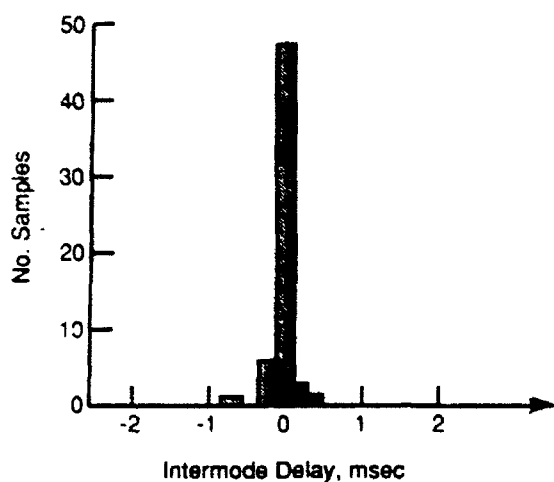
was estimated from ten consecutive cross power spectra. Thus, each estimate of time delay and transmitter location was based on 0.354 second of signal observation time.

Histograms of the interpath delay estimates and of the relative amplitude of surface wave to skywave are shown in Figure 2. Each of the histograms is composed of 64 samples and relative frequency of occurrence is indicated with respect to the sample mean value. The interpath delay histogram of Figure 2(a) indicates that 75% of the data are contained within  $\pm 0.125$  msec of the sample mean of 1.997 msec. The amplitude histogram of Figure 2(b) shows that 65% of the data are contained within  $\pm 0.1$  of the average value of 1.57. The amplitude ratio indicates that the surface wave contains approximately 4 db more power than the skywave.

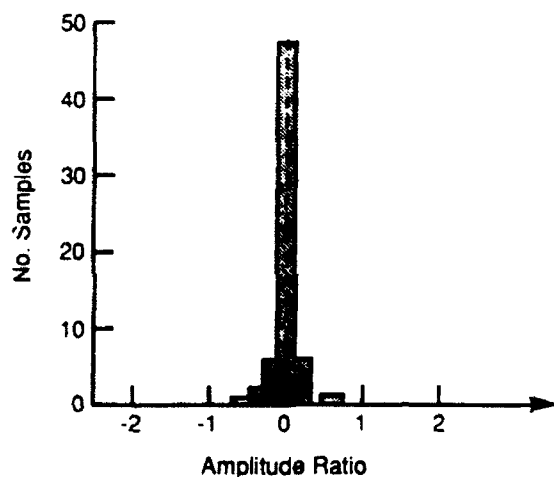
In the application of the normalized MUSIC cepstrum, it became apparent that the technique produced a peak value for *differential* intersensor delay and did not produce a unique estimate for each intersensor delay. In particular, consider Equation (4b) and the terms involving  $\tau_g$  and  $\tau_s$ , which represent group delay between antennas for the groundwave and the skywave, respectively. Factoring out a common term relating to  $\tau_g$ , we obtain

$$A_2 = [A_g + A_s e^{-\omega(\Delta\tau + \tau_m)}] e^{-\omega\tau_g} \quad (11)$$

where  $\Delta\tau = \tau_s - \tau_g$ . For narrow bandwidths and relatively short spacing between antennas, the term



(a) Scatter Relative to Mean Value of 1.997 msec



(b) Scatter Relative to Mean Amplitude Ratio of 1.57

Figure 2. Histograms of Estimated Interpath Delay and Amplitude Ratio

$\exp(-j\omega\tau_g)$  is essentially a constant. For this reason, the normalized MUSIC cepstrum produces a sharp peak for  $\Delta\tau$  and arbitrary  $\tau_g$ ; therefore, in our solution process, we set  $\tau_g = 0$  and estimate  $\Delta\tau$ , the differential intersensor delay. It is worthwhile to note that the normalized MUSIC cepstrum would lead to a unique estimate of  $\tau_g$  and  $\tau_s$  for large bandwidth signals, as for example spread spectrum formats. In this case, the term  $\exp(-j\omega\tau_g)$  would exhibit a periodicity which could be estimated. An important observation is that the normalized MUSIC cepstrum appears to be insensitive to systematic phase offsets between receiver channels. The interpath delay and differential intersensor delays produce periodicities which may have a constant offset (or trend), and this does not affect the MUSIC cepstral estimate.

Histograms showing the distribution of differential intersensor delay estimates for the four baselines are plotted in Figure 3. The baseline annotations refer to the antenna numbers indicated in Figure 1. The least robust scatter is exhibited by the histogram of Figure 3(a), which represents data acquired on the 2-3 baseline aligned nearly broadside to the incident wavefield. Because the aperture is aligned with the wave interference field, this baseline was expected to evidence a lower degree of statistical stability. The differential intersensor delay estimates for baselines 1-2 and 2-4, shown in Figures 3(b) and 3(c), are somewhat more robust. This was an expected result since the baselines are more nearly directed along the propagation path of both signals, and as a consequence they have a greater sampling aperture relative to the wave interference field. The histograms shown in Figure 3(a)-3(c) evidence a bimodal characteristic, and it is presently not clear to us what mechanisms cause this to occur. The most robust scatter of data is shown in Figure 3(d) for the 1-4 baseline which was oriented along the direction of propagation and as a consequence, had greatest sampling aperture relative to the wave interference field.

The differential intersensor delay times and the interpath delay time can be used in concert to estimate the location of the transmitter. Since this is being done from a single receiving site and all the measurement parameters are derived from the received signal, this process is referred to as passive SSL. A scatter plot of location estimates is shown in Figure 4. Each location estimate was based on data acquired over an interval of 0.354 second, and the entire collection of fixes were taken in contiguous intervals over a period of 5.31 seconds. The receiving site on the Texas coast line is shown by the square and the transmitting site in the Gulf of Mexico is shown by the large "X". The estimated locations are indicated as triangles. The scatter of estimated locations appear to evidence a systematic bias south of the true transmitter position.

To quantify the bias in the location estimates, we computed azimuth and range error for each point. Histograms of azimuth and range error are shown in Figure 5. Each plot is based on 16 location estimates. The azimuth error histogram of Figure 5(a) indicates a bias toward the south of 8.6 degrees, with the greatest concentration of data located between zero and twenty degrees error. The histogram of range error shown in Figure 5(b) indicates a systematic bias of 18.5 Km. These data are somewhat more uniformly scattered than the azimuth error data. Averaging the position estimates results in a miss distance of 39.6 Km or 17% of the true range of 225 Km.

#### 4. DISCUSSION

In the mathematical development of Section 2, it was assumed that cross correlation terms involving the transmitted signal and the additive noise were negli-

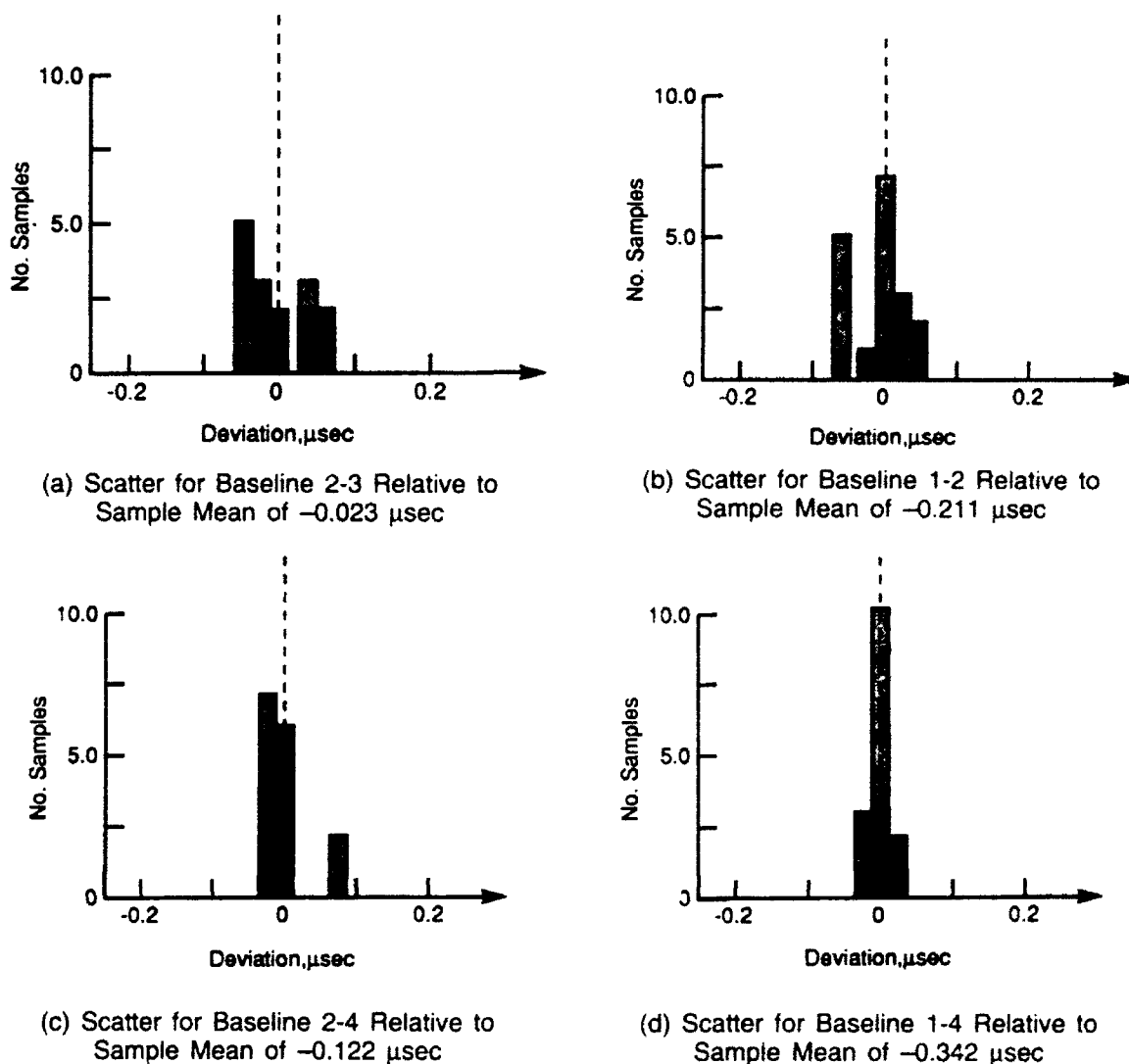


Figure 3. Histograms of Estimated Differential Intersensor Delay,  $\Delta\tau$

gible. The implication of this assumption is that the expected phase error in Equation (8) is zero. This will not be the case if the additive noise terms are mutually coherent with each other or with the transmitted signal, then the complex argument of Equation (8) will be corrupted and one would expect to observe an attendant degradation in the accuracy of delay estimates. Also we have implicitly assumed that the sampling process is statistically wide sense stationary and that a time average is equivalent to an ensemble average; therefore, if the sampling duration is too short, the stationary condition may not hold.

A second simplifying assumption made in the development of the algorithm was that the antennas had ideal omnidirectional response. For the array deployed in this work, the crossed loop antennas satisfied this condition. In the more general case, if the antennas used have directional response patterns or if the patterns are perturbed by nearby reradiating struc-

tures, then these effects must be taken into account in the array steering vector of Equation (10).

Aside from algorithmic issues, factors which may affect the performance of the technique are signal bandwidth, antenna separation and relative mode amplitude. The issue of signal bandwidth is important since it will determine the number of periods which are to be used to estimate the interpath delay time. From Equation (5c) it is observed that the phase function has a spectral period of  $1/\tau_m$ . Thus, if there were 2 msec delay between modes, then we would expect to observe a periodicity of 500Hz in the complex argument of the coherence function, and a signal bandwidth of 2Khz would support four periods. On the other hand, if the signal bandwidth is very narrow, say 200Hz, then a full period is not present and the technique may not provide a satisfactory estimate.

The separation between antennas is important since it will determine the magnitude of the periodicity in

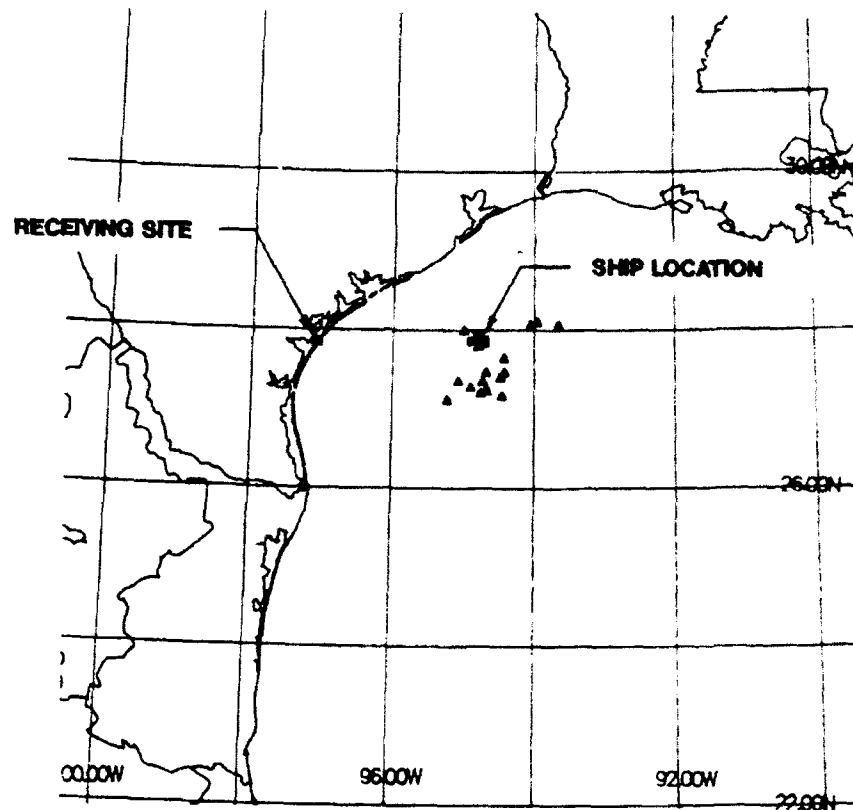
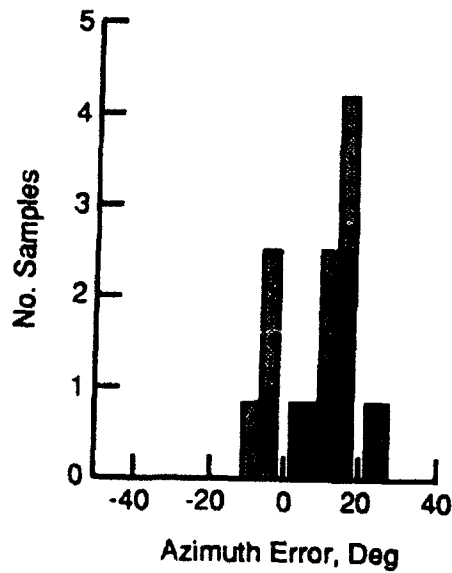
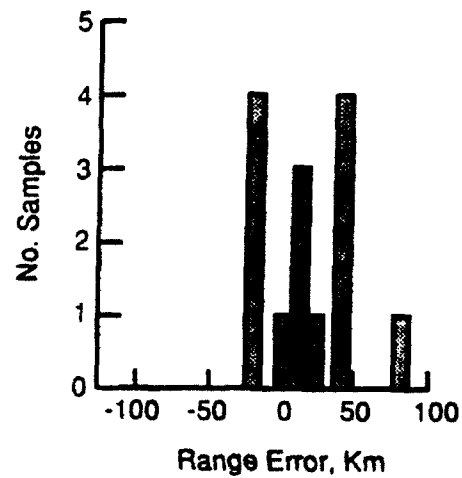


Figure 4. Scatter Plot of Estimated Transmitter Locations



(a) Azimuth Error Relative to True Value of 89.1 Degrees



(b) Range Error Relative to True Value of 225 km

Figure 5. Histograms of Error in Azimuth and Range Estimates

the phase function of Equation (5c). This can be seen from the argument of the sine in the numerator which involves  $\tau_g$  and  $\tau_s$ . As the separation between antennas decreases, the intersensor delays approach zero and the sine term consequently becomes small. To maximize the sine term and at the same time minimize the denominator, we would prefer to set the distance between antennas such that  $\tau_g + \tau_s = (2k + 1)\pi$ . Obviously, this is difficult to do since it requires knowledge of the AOA for each signal.

The relative path amplitudes have a similar effect on the amplitude of the periodicity of the phase function. It is clear from Equation (5c), that as the path amplitude terms  $A_g$  and  $A_s$  decrease, the magnitude of the phase ripple decreases. An area for further investigation is to determine the range of relative amplitudes between paths for which this approach is effective.

The experimental proof-of-concept study done in this effort was particularly encouraging since it indicated that under reasonably favorable conditions, the passive SSL technique would work. The location data of Figure 4 evidenced a bias toward the southeast which we believe may have been due in large part to the unfavorable antenna array geometry shown in Figure 1. To obtain more nearly optimum spatial sampling, the angle formed at antenna #1 should have been 60 degrees as opposed to 35.5 degrees. Future efforts will concentrate mainly on delta array geometry to provide improved spatial sampling of the signal interference pattern.

One of the major sources of error in conventional SSL technology is the identification of the propagation path between the transmitting and receiving sites. In the case of HF communication, one must consider the possibility of  $E$ ,  $F_1$  or  $F_2$  layer reflections in the ionosphere. If we restrict consideration to one hop multipath, the passive SSL technique implicitly resolves the issue of mode identification and is not affected by this source of error. At longer ranges where the propagation paths may consist of multiple hops, both conventional and passive SSL techniques will require a more sophisticated propagation model for

accurate location estimates. However, for short or intermediate range fixing, the passive SSL approach offers the promise of significant improvement in reducing location error due to incorrect identification of propagation modes.

## REFERENCES

1. Gething, P.J.D., "Radio Direction Finding and Superresolution", 2nd Ed., London, UK, Peter Peregrinus Ltd., 1991.
2. Warrington, E.M. and Jones, T.B., "Measurements of the Direction of Arrival of HF Skywave Signals and Single Site Location Using a Seven Element Wide Aperture (294 m) Interferometer Array", IEE Proceedings - Pt. F, 138, 2, April 1991, pp121-130.
3. Horing, H.C., "Comparison of the Fixing Accuracy of Single-Station Locators and Triangulation Systems Assuming Ideal Shortwave Propagation in the Ionosphere", IEE Proceedings - Pt. F, 137, 3, June 1990, pp173-176.
4. Sherrill, W.M., Travers, D.N. and Martin, P.E., "Advanced Radiolocation Interferometry", in "Proceedings of the Fourth Allerton House Conference on HF Radio Direction Finding and Radiolocation Research", Urbana, IL, June 1971.
5. Carter, G.C., Nuttall, A.H. and Cable, P.G., "The Smoothed Coherence Transform", Proceedings of the IEEE, 61, 10, October 1973, pp1497-1498.
6. Davies, K., "Ionospheric Radio", London, UK, Peter Peregrinus Ltd., 1990.
7. Rendas, M.J.D. and Moura, J.M.F., "Cramer-Rao Bound for Location Systems in Multipath Environments", IEEE Transactions on Signal Processing, 39, 12, December 1991, pp2593-2610.
8. Schmidt, R.O., "Multiple Emitter Location and Signal Parameter Estimation", IEEE Transactions on Antennas and Propagation, AP-34, 3, March 1986, pp276-280.

## DISCUSSION

### L. BERTEL

Do you use the loop antennas independently or as a polarization filter?

### AUTHOR'S REPLY

The loops are quadrature summed to discriminate between the X and O modes. The E-W loops are retarded 90° and summed with the N-S loops to obtain quad left and reject the X mode. Similarly, the N-S loops are retarded 90° and summed with the E-W loops to reject the O mode.

### G. HAGN

Could you comment on the use of the MUSIC algorithm fed by your crossed loops when used on short path skywave signals? It should be possible to use data from the MUSIC algorithm to generate measured ionograms in the standard form, but with azimuth and elevation information by mode as well. A ray-tracing propagation prediction model (e.g., AMBCOM) could be used to generate synthetic ionograms for comparison with the measured data, and the input ionosphere varied to obtain a best fit to not only the ionogram in its standard form but also the elevation-plane angle of arrival. The adjusted ionosphere (giving the best fit) could then be used for backward ray tracing of the MUSIC-derived signals to obtain the estimate of emitter location. But this is the subject for further research.

### AUTHOR'S REPLY

The essential fixing procedure we have used is based on the intersensor and interpath delays and employing the Briet and Tuve theorem. It is highly probable that a high ray-low ray combination does not fit this model or the near-vertical incidence case of ground wave and tilted ionospheric skywave. It would appear that we need to do research on our understanding of more complicated ray paths and implement this into our fixing model. The approach you suggest of iteratively looking for agreement between our empirical mode decomposition measurements and theoretical models seems to be a reasonable starting point.



## PROCEDURES FOR DETERMINING THE LOCATION OF AN HF RADIO TRANSMITTER

Leo F McNamara  
ANDREW SciComm Inc  
2908 National Drive  
Garland, Texas 75041, U.S.A.

### 1. SUMMARY.

The location of an HF transmitter may be derived by either of two methods, depending on what observations of the target transmitter are available. In the traditional approach currently adopted for strategic DF networks, the location is determined by triangulation using Lines of Bearing (LOB) of the signals observed at two or more HF-DF sites. This approach is called here the LOB approach. Alternatively, the location may be derived by observations at a single DF site, if the observations are extended to include that of the elevation angle as well as the bearing of the incoming radio waves. Triangulation in the vertical plane containing the observed LOB, using a model of the propagation medium (the ionosphere), leads to an estimate of the position of the target transmitter. This approach is known as the Single Site Location (SSL) method. The success of the LOB approach depends heavily on two or more DF sites being able to hear and process the signals from the target transmitter, as well as on having reasonably favorable geometries. When these conditions are not met, the LOB approach fails. The SSL method, on the other hand, is applicable under almost all conditions. This paper therefore compares and contrasts the two approaches to determining the location of an HF transmitter. The relative advantages of the SSL approach are shown to be such that all HF-DF sites should logically be SSL sites.

### 2. INTRODUCTION

With a few little-known exceptions in the last three decades, but more in recent years, the location of an HF transmitter has been determined by a process of triangulation, using measured Lines of Bearing (LOB) or azimuths of the signals received at two or more HF Direction Finding (DF) sites. Ideally, at least three DF sites should be able to observe the target transmitter. In an alternative approach which is becoming more widely adopted, the DF system also measures the elevation angle of the incoming radio waves. Given a knowledge of the ionosphere along the path to the target, it is then possible to determine the point of origin of the signals, and hence the location of the target, by simulating the propagation of radio waves back along the observed direction. This procedure requires observations from only one DF site, and is accordingly known as the Single Site Location (SSL) technique. For increased accuracy, the procedure also requires real-time observations of the ionosphere, using an ionosonde at the SSL site.

SSL systems have been deployed operationally for at least 30 years in situations for which the traditional LOB approach is not possible, either because the circuits were so short that the measured LOBs were completely unreliable indicators of the true bearing to the target, or

because only one DF site was available in the first place. In recent years, the closure of overseas bases for both political and economic reasons has pointed to the need to get by with fewer assets, and has highlighted the advantages of being able to determine the location of an HF transmitter using just one DF (actually SSL) site.

This paper compares the traditional LOB or network approach with the SSL approach, discussing the strengths and weaknesses of the two approaches. Generally speaking, there are some areas in which one approach is clearly superior to the other, and some areas where there is not much to choose between the two. There is still some room for improvement in the LOB approach, using modern high-resolution systems which can discriminate between the LOBs for different propagation modes, together with propagation simulation, to isolate subsets of the observations which are likely to yield the most accurate estimates of the position of the transmitter (position estimates, or PE). In spite of recent advances in our understanding of the ionosphere and how it affects the measured angles of arrival (AOA; i.e., azimuth and elevation) on different circuits, there is still much to be learned about the SSL technique and how it may best be applied. It seems highly likely, however, that the DF systems of the future will all be SSL systems, because of the overwhelming advantages of the SSL approach.

The five sets of simulation results presented here all correspond to realistic situations, although the exact locations of the networks are not given. The results do depend, of course, on the actual locations because of the variations of the ionosphere with latitude and longitude.

### 3. METHODS OF POSITION ESTIMATION

#### 3.1 Introduction

Within the SSL approach, three different methods of position estimation are used, depending on the length of the circuit,  $R$ , and the observed elevation angle,  $\beta$ :

Short Ranges	$\beta > 75^\circ$	$R < 100$
Classical SSL with Tilt Correction		
Medium Ranges	$\beta > 50^\circ$	$R > 150$
Classical SSL	$\beta < 75^\circ$	$R < 600$
Long Ranges	$\beta < 50^\circ$	$R > 600$
SMART-SSL™		

The transition angles and ranges,  $50^\circ$ ,  $75^\circ$ , 150 km and 600 km, are all somewhat arbitrary, since they depend on the state of the ionosphere at the time, and on the relative

accuracies of the three methods of position estimation. SMART-SSL™ is an acronym for the Synoptic Modelling And Ray Tracing approach to SSL. The Classical SSL method relies on some simple theorems of radio wave propagation, as described in Section 3.2. Reviews of the SSL technique have been given in references [1] and [2].

The SSL approach is most accurate at medium ranges, say 200 to 600 km, since the effects of tilts in the ionosphere at the reflection point are small, and the local ionogram is a good indicator of what the ionosphere is like at the circuit reflection point. Application of the SSL approach is more difficult at short ranges, for which the calculated position estimates depend strongly on the tilts existing at the reflection point, and at long ranges, for which observations of the ionosphere at the SSL site may be a poor indicator of the ionosphere at the remote reflection point.

### 3.2 The Classical SSL Method of Position Estimation

The Classical SSL method is the most broadly applicable and commonly used of the various methods of range estimation. Given a local vertical incidence (VI) ionogram, it is simple and quick to apply, and can be very accurate. The Classical SSL method assumes that the actual propagation may be modelled by assuming that reflection takes place from a simple horizontal mirror at the appropriate height. If the Earth were flat, and if the radio signals were reflected by a horizontal sporadic-E (E<sub>s</sub>) layer, for example, the ground range TR would be related to the measured elevation angle,  $\beta$ , by

$$TR = 2h / \tan \beta.$$

The Classical SSL method is a generalization of this idea, relying on the availability of a suitable VI ionogram to provide the height  $h$ , and on the applicability of some fundamental laws of radio propagation through the ionosphere.

The method is based on three simple fundamental relations - the Secant Law, Breit and Tuve's Theorem and Martyn's Equivalent Path Theorem [3]. If  $f$  is the frequency which is reflected at normal incidence from the same real height as the operating frequency,  $f$ , the Secant Law states that

$$f = f_0 \sec \chi,$$

where  $\chi$  is the angle of incidence of the ray at the base of the ionosphere. For a flat Earth,  $\chi$  is the complement of the elevation angle. The frequency  $f_0$  is called the "Equivalent Vertical Incidence Frequency".

Application of the so-called Classical SSL Method of range (or position) estimation proceeds as follows :-

1. The incoming signal at the frequency  $f$  is observed to have an elevation angle  $\beta$ .
2. The secant law is used to calculate the equivalent vertical frequency,  $f_0$ .
3. The virtual height  $h'(f_0)$  is obtained from a local VI ionogram, at the frequency  $f_0$ .
4. The signal is assumed to have travelled a triangular path, with an elevation angle  $\beta$ , and with reflection occurring at an altitude equal to  $h'(f_0)$ .
5. The range to the transmitter, TR, is then simply  $2h'/\tan \beta$ .

The extension to a curved Earth is straight forward. The Classical method ignores the Earth's magnetic field when calculating the equivalent vertical frequency, but takes it into account when the value of  $h'(f_0)$  is scaled from the ionogram.

### 3.3 Short-Range Tilt Correction

In general, tilts play a much greater role on shorter circuits. The effects are often so large, as seen in widely varying azimuths, that the standard LOB approach is not applicable. It is instructive to consider the % error in the calculated range,  $D$ , due to a tilt  $\theta$  in the direction of propagation on a short circuit with elevation angle  $\beta$ . We find

$$\delta D/D = -1.75\theta/\sin\beta/\cos\beta \%,$$

where  $\theta$  is in degrees. Thus if  $\beta = 85^\circ$ , corresponding to an F-mode ground range of around 50 km, the error is 20% of range per degree of tilt. The error drops to 5% for  $\beta = 70^\circ$ , corresponding to a range of about 200 km. The transverse component  $\theta$ , of the tilt deviates the ray path out of the great circle and produces a bearing error. If reflection is attributed to a plane mirror at a height  $h$  above a plane earth, then the bearing error is approximately [4]

$$\delta\phi = (2h\theta) / D.$$

Taking a reflection height of 300 km yields bearing errors of 12 and 3 times the transverse tilt angle for circuit lengths of 50 and 200 km respectively. The effects of any tilts are thus most important when  $\beta$  is large, or  $D$  is small i.e. for short circuits.

In order to correct position estimates for the effects of tilts, real-time AOA observations are needed from a known transmitter in the same general area as the unknown transmitter. These are used to derive estimates of the ionospheric tilts at the circuit reflection point, assuming a tilted mirror at the height  $h'(f_0)$ . The tilts so derived are then used in conjunction with the unknown transmitter AOA observations to find the range to that transmitter, along a corrected azimuth. Any other HF transmitter outside the range of the ground wave may be used as a remote check target. An alternative approach is to use an appropriate ionosonde located at the SSL site to provide observations of the tilts in the local ionosphere.

The possibility of tilt correction arises only for an SSL system working short-range targets. The technique cannot be applied when only the LOB is measured, since elevation angles of the target and check target signals, as well as the virtual height  $h'(f_0)$ , are also required.

### 3.4 Long-Range Position Estimation

Long-range position estimation techniques are invoked at ranges beyond the validity of the Classical SSL method of position estimation i.e. when the ionogram obtained at the SSL site is not expected to be a good indicator of what the ionosphere is like at the midpoint of the long circuit. As a general rule of thumb, the validity of the local ionogram starts to deteriorate at a range equal to the height of the reflecting layer - 100 km for the E and E<sub>s</sub> layers, and 300 km for the F layer. Beyond these distances, it is a question of how the correlation decreases with the separation of the SSL site and the circuit midpoint. Obviously the correlation decreases more rapidly at sunrise and sunset, and in the presence of other large-scale horizontal gradients.

The technique adopted for long-range position estimation is basically one of ray re-tracing through a model of the ionosphere which has been updated in real time to match the characteristics scaled from the ionogram obtained at the SSL site. Radio waves are traced in simulation through the model, in the opposite direction to that at which they were observed to arrive at the SSL site, and the point at which the ray hits the ground is the calculated location of the transmitter.

The essential elements in this ray re-tracing procedure are

1. A reliable synoptic model of the ionosphere which can be updated in real time using parameters derived from the local ionogram
2. A method of raytracing which has an accuracy well exceeding the accuracy with which the ionosphere can be specified along the whole circuit.

### 3.4.1 Ionospheric Modelling Techniques

We refer here only to techniques which are valid for mid-latitude circuits, since the bulk of the AOA data currently available for long-range circuits is from mid-latitude sites. The  $N(h)$  profile at mid latitudes can be fairly well defined in terms of 10 parameters - the critical frequency, height of maximum electron density and parabolic semithickness for each layer, and the E-F valley width :-

1. E layer -  $f_oE$ ,  $h_mE$  and  $y_mE$
2. E-F valley - width
3.  $F_1$  layer -  $f_oF_1$ ,  $h_mF_1$  and  $y_mF_1$
4.  $F_2$  layer -  $f_oF_2$ ,  $h_mF_2$  and  $y_mF_2$ .

World maps or models exist for these parameters, or of parameters such as  $M(3000)F_2$  from which they may be derived. The least well known seem to be  $y_mF_2$ ,  $h_mF_1$  and  $y_mF_1$ . Any convenient profile may be erected at a given point along a circuit, provided it matches the modelled values of the 10 parameters, since there are currently no definitive models of the profile shape between the various anchor points.

Updating of the synoptic model is most simply performed via an ionospheric index, or effective sunspot number, which forces the model of the various parameters to match the values observed by the ionosonde at the SSL site. A different effective sunspot number may be used for each of the characteristics defining the  $N(h)$  profile. The index derived from  $f_oF_2$  is called the T index, since the maps of  $f_oF_2$  which are used are given in terms of this index [5]. In general, as much relevant information as possible should be derived from the local VI ionogram, since the ultimate accuracy of the calculated transmitter locations depends to a large extent on how well the local ionogram can be used to characterize the ionosphere along the whole circuit.

### 3.4.2 Ray Re-Tracing Techniques

Given a sufficiently accurate model of the ionosphere along the circuit, the location of the unknown transmitter can be determined by raytracing through that model back along the direction of the incoming wavefront. The raytracing may be approached at various levels of complexity, depending on the relative computation times and the accuracy with which the ionospheric model can be specified. If the profile shapes are quasi-parabolic, solutions for the range are given in closed form in terms of the profile characteristics, the wave frequency and the AOA. See, for example, reference [6]. For arbitrary profile shapes, the Jones-Stephenson [7] three-dimensional

raytracing program is the method of choice.

Some of the possible approaches are:-

1. Analytic two-dimensional raytracing through a multi-segmented quasi-parabolic  $N(h)$  profile with no variation along the circuit, neglecting the Earth's magnetic field.
2. Analytic two-dimensional raytracing through a multi-segmented quasi-parabolic  $N(h)$  profile with the ionospheric characteristics being allowed to vary along the circuit, neglecting the Earth's magnetic field.
3. Numerical three-dimensional raytracing through an arbitrary  $N(h)$  profile which varies along the circuit, including the Earth's magnetic field.

All three methods were used by McNamara [8], who found that in general it is not possible to specify the ionosphere along the circuit well enough to cause the more complex methods (2 and 3) to be preferred over the simplest one (1).

The errors in position estimates are usually described in terms of the "miss distance", which is the distance between the calculated and correct locations of the transmitter. For long-range transmitters, the miss distances are made up of contributions from various sources [9], the relative importance of which will depend on the particular circumstances :-

1. A bearing error due to the lack of observations over the full period of any travelling ionospheric disturbance (TID).
2. Biases in the peaks of the elevation angle distribution for multi-moded propagation, due to fading and data filtering.
3. A bearing error due to a large-scale synoptic gradient not accurately represented in the ionospheric model, or not allowed for in the raytracing.
4. A range error due to differences between the model and actual  $N(h)$  profiles along the circuit.
5. Range and bearing errors due to errors in, or limitations of, the raytracing algorithms.
6. Use of the wrong order of the propagation mode. The observed AOA do not seem to contain any information which could be used to identify uniquely the number of hops that the signals have made on their way from the transmitter.
7. Confusion caused by  $E_s$ . As with (6), the observed AOA do not indicate whether or not  $E_s$  has been involved in any way.

In principle, the raytracing procedures may be developed to any required level of accuracy, so the major source of errors in the miss distances will be our inability to specify the ionosphere accurately along the whole circuit. In a strict SSL approach, the only ionospheric knowledge available is that obtained at the SSL site itself, so it will not usually be possible to specify the ionosphere along the circuit with arbitrary precision.

## 4. THE PRODUCTIVITY OF AN HF-DF NETWORK

### 4.1 Introduction

Several important questions arise when a network of HF-DF sites is proposed:

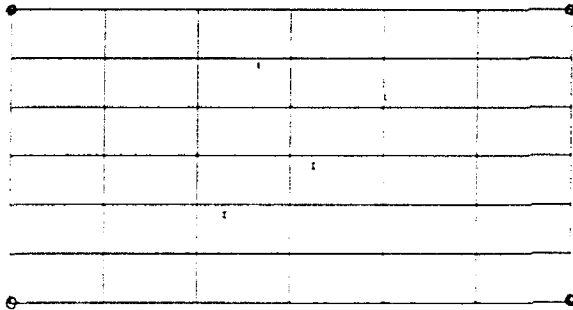
1. How many observing sites are needed to cover the area occupied by the target transmitters?
2. Where should each of the sites be located for optimum

coverage of the target transmitters?

3. What are the advantages to be gained by making each site an SSL site, rather than just an LOB site?

4. What are the advantages arising from the use of a system with a higher sensitivity?

Questions 3 and 4 are addressed here by using HF propagation simulations, with special reference to a physically large DF network (Network A) illustrated in Figure 1. Ideally, questions 1 and 2 should be addressed in the early planning stages, using techniques similar to those described here. However, in some applications, the sites and their dispositions are dictated by external needs, such as the availability of logistic support.



1. Locations of the 4 DF sites (X) and the limits (O) of the grid of transmitters considered for Network A. The latitude and longitude lines are given for every 10°

It is important to be able to estimate both the probability that the network can determine the location of a given transmitter, and the accuracy of the calculated location. This section addresses the former question; the latter will be addressed in the next section. At present, we will just note that any simulations which portray the accuracies that can be expected from a DF network can by themselves be misleading, since the calculated accuracies can be realized in practice only if the propagation conditions and system sensitivities are such that all the necessary members of the network can hear a given transmitter.

#### 4.2 Method

The ionosphere and its variations are so complex that it is virtually impossible to estimate a priori just how often one or more sites can work the same transmitter, and thus how often the SSL approach would be required, or how often multi-site fixes could be based on observations from more than two sites. This problem has therefore been simulated, using the standard HF propagation prediction program IONCAP [10]. The DF sites for Network A are very widely separated, covering 3000 km in latitude.

Propagation predictions have been made using IONCAP (Method 17; Condensed System Performance) for circuits between the four sites and transmitters located at 7 x 5 gridpoints (7 latitudes, 10° apart; 5 longitudes, 15° apart). The area to be monitored is illustrated in Figure 1. For Network A, predictions were made for 4 sites; 7 latitudes; 5 longitudes; 3 months; 2 sunspot numbers; 8 hours; 11 frequencies. There were  $7 \times 5 \times 3 \times 2 \times 8 \times 11 = 18480$  calculations of the S/N ratio made for each site. Each calculation is considered here as the calculation for an effective transmitter, and there are 18480 of these for each DF site.

The IONCAP runs were made for

1. Constant gain antennas
2. A minimum elevation angle of 5°
3. 1 kW transmitter power
4. 1 Hz bandwidth
5. Rural-level man-made noise
6. Antenna gains of 10 dB.

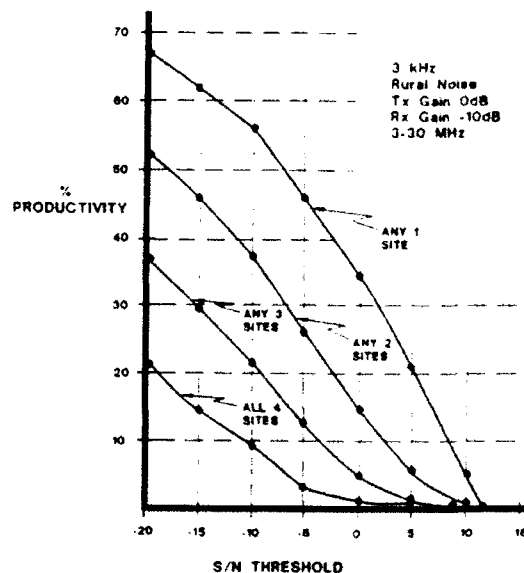
There is no simple way to adjust the calculated values of S/N to alter the restrictions 1 and 2. However it is a simple matter to reprocess the data to allow for different values for lines 3 to 6.

Interrogation of the output files allows the calculation of the number of transmitters for which the calculated S/N ratio exceeds a specified threshold. The ratio of this number to the total possible number (18480) is called the **productivity** of the DF network. Threshold values of -20 to 20 dB in steps of 5 dB have been considered. The analysis can be performed for different values for lines 3 to 6. The case considered here is for 100 watt transmitter powers; a 3 kHz bandwidth; rural-level man-made noise; 0 dB transmitter antenna gain; -10 dB receiver antenna gain. These are typical values for DF work.

#### 4.3 Results for Network A

##### 4.3.1 Productivities for N-site Fixes

Figure 2 shows the percentage of cases for which the calculated S/N ratio for the 18480 effective transmitters was above a specified threshold value at one or more DF sites. The curve labelled "Any 1 site" is the productivity of the network when only one site has to be able to "hear" (i.e., monitor and process) each transmitter, as in the SSL approach. The curve labelled "Any 2 sites" is for the case when any two or more of the four sites can hear a transmitter; the curve labelled "Any 3 sites" is for the case when any three or more of the four sites can hear a transmitter; the curve labelled "All 4 sites" is for the cases in which all four of the DF sites can hear the one transmitter.



2. Productivity of Network A, as a function of system threshold, for observations from different numbers of DF sites, assuming typical values for the various parameters

Suppose, for example, that the DF system installed at each site can work at 0 dB S/N ratio. Then 34.5% of all the transmitters could be worked by at least one of the four systems, if they were SSL systems. If any 2, any 3 or all 4 sites in the network are required to be able to observe a transmitter, the productivities are 15.0, 5.4 and 0.8%.

If an extra 10 dB can be provided by higher-gain antennas, or from 1000 watt transmitters, for example, the chances of any one of the four SSL systems being able to work a transmitter increases from 34.5% to 56.1%. An increase of power has a larger effect when any two sites are used in the LOB approach, the productivity almost trebling from 15.0 to 37.6%. This means that the LOB approach relies more heavily on high S/N levels than does the SSL approach.

#### 4.3.2 Productivity Advantage of the SSL Approach

Consider again the results given in Figure 2, for a 0 dB S/N threshold. Relative to the 1-site (SSL) value, the productivities when 2, 3, or 4 LOBs are required are 1 : 0.43 : 0.16 : 0. In other words, if the 4 sites are equipped with SSL equipment, they can work  $1/0.43 = 2.3$  times as many transmitters as four simple LOB sites if two LOBs are required (yielding a "cut"); or  $1.0/0.16 = 6.25$  times as many transmitters if three LOBs are required (yielding a "fix"). The probability of all 4 sites being able to hear the given transmitters at a 0 dB S/N ratio is effectively zero. Similar scenarios can be deduced for any other threshold value.

It is also interesting to consider how many of the  $3 \times 2 \times 8 \times 11 = 508$  nominal transmitters placed at each gridpoint may be heard from one or more DF sites. As expected, the productivity is highest for those gridpoints contained within the network, and closest to the four DF sites. The value of the SSL approach is especially great for transmitter locations remote from the DF sites. For example, transmitters along a line at the extreme west of the area being monitored can never be heard simultaneously by three or four DF sites with a threshold of 0 dB, and only on 2% of occasions can they be heard by two DF sites. However at a latitude close to that of the southern-most DF site for example, 27% of transmitters could be heard from at least one SSL site.

#### 4.3.3 Advantages of a Lower S/N Threshold

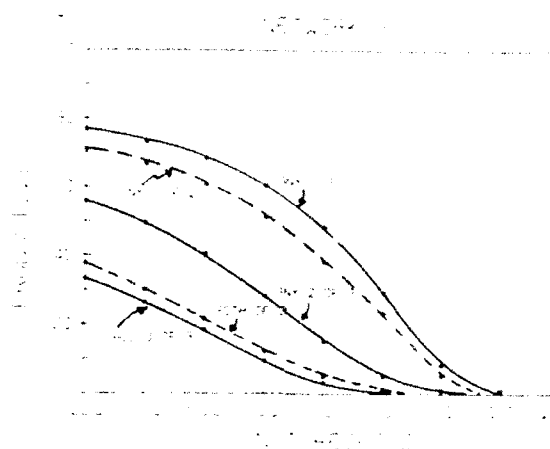
Figure 2 may also be used to derive the advantages offered by an SSL system which can work at a lower S/N ratio than another. For example, curve 1 of Figure 2 shows that SSL systems with a threshold of -5 dB can work  $46 / 21 = 2.2$  times as many transmitters as a system with a threshold of +5 dB.

If SSL systems with a threshold of -5 dB were used at the four sites, in place of LOB systems with a threshold of +5 dB, and at least two LOB measurements are required, approximately  $2.2 \times 1.8 = 4$  times as many 100 watt transmitters distributed around Network A could be worked.

#### 4.4 Network B

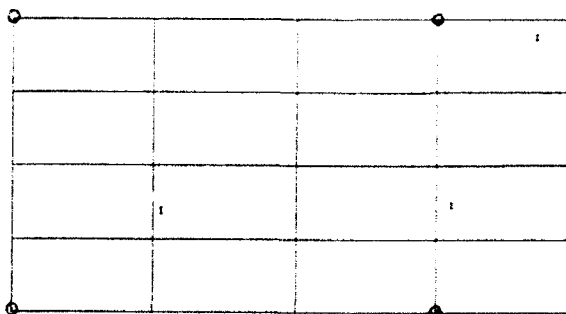
The continuing closure of overseas sites has highlighted the precarious nature of the current strategic DF networks. The solid curves in Figure 3 are for the case of a three-station net, monitoring transmitters at the  $7 \times 9$  gridpoints at pseudo-latitudes  $0^\circ\text{N}(5)40^\circ\text{N}$  and pseudo-longitudes  $0^\circ\text{E}(5)30^\circ\text{E}$  (Figure 4). It is required that (top curve) any one of the three stations be able to hear a given

transmitter (the SSL solution); that any two stations be able to hear a transmitter (the LOB solution, but relying on a cut); and that (bottom curve) all three stations be able to hear a transmitter (the LOB solution, demanding a fix, rather than a cut). Taking a -5 dB threshold as typical, it can be seen that the productivity drops from 48.7% to 15.3% to 5.3%.



3 Productivity of Network B, as a function of system threshold, for observations from different numbers of DF sites, assuming typical values for the various parameters. The full lines correspond to the availability of all three DF sites, while the dashed lines correspond to the case in which DF2 is not available.

The dashed curves in Figure 3 correspond to a two-station net, DF1 and DF3. The differences in productivity between the two sets of curves graphically illustrate the decrease in productivity caused by the removal of DF2 from the net. To start with, it is no longer possible to obtain a fix, so the only comparison which can be made is between the productivities of the SSL and LOB approaches relying on a cut. At 0 dB, the relative productivities are 39.0 and 5.3%, giving the SSL approach a seven-fold advantage in productivity. For the three-station net, the advantage was  $48.7 / 15.3 = 3.2$ , so the loss of the third station has doubled the advantage offered by the SSL approach.



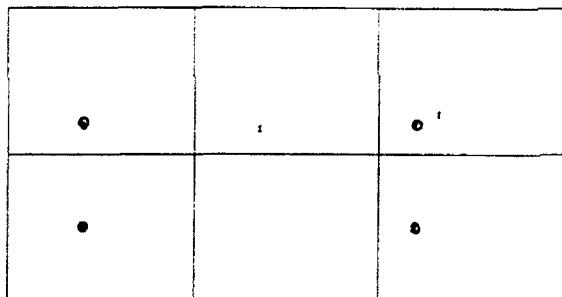
4. Locations of the 3 DF sites (X) and the limits (O) of the grid of transmitters considered for Network B. The latitude and longitude lines are given for every  $10^\circ$ .

The productivity of an SSL network drops by only 20% when DF2 is removed. The effect of removing DF2 from a three-station LOB net is to drop the productivity from 15.3% to 5.3%, which is a factor of three. Networks of SSL stations are thus much more stable against the loss of one station than are LOB nets.

The loss of DF2 as an LOB site has most effect on those gridpoints to the south of that site. For example, the productivity at the pseudo-point (0°N, 20°E) drops from 26% to 2%. Exactly how the productivity changes is a complicated function of the frequencies, times of day etc which are considered. The productivity for that gridpoint would have fallen only down to 39% (from 49%) if DF1 and DF3 had been SSL sites.

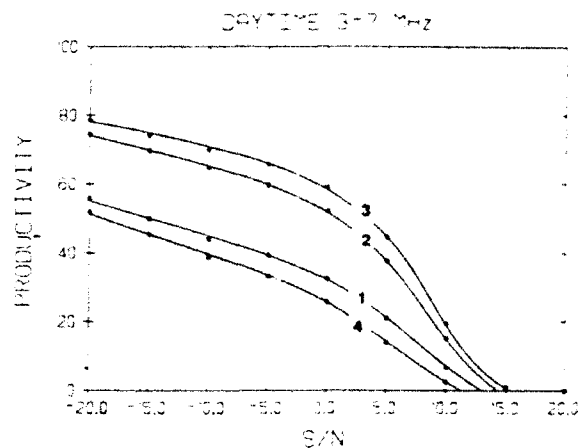
#### 4.5 Network C

This section considers the problem of determining the locations of HF transmitters operating at the extreme edge of the area being monitored, as illustrated in Figure 5. An array of HF transmitters is placed (in simulation) at gridpoints within a rectangle covering pseudo-latitudes of 0°N to 7°N, and pseudo-longitudes of 0°E to 18°E. The productivity results have been grouped into day (12 to 23 UT), and night (00 to 11 UT), and separate calculations have been made for frequency ranges of 3 to 7 MHz and for 3 to 16 MHz, to cover two possible scenarios for the assumed transmitters. The results for the lower frequencies are given in Figures 6 (day) and 7 (night).

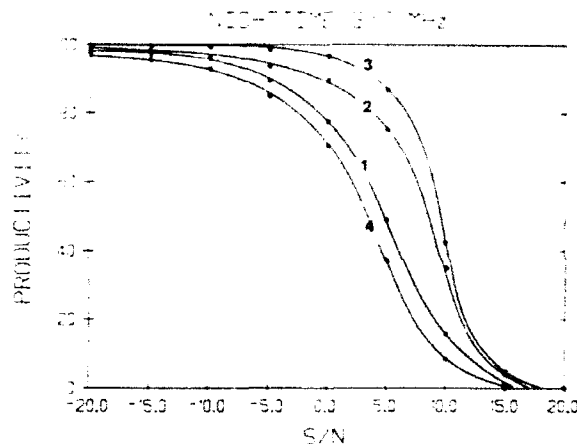


5. Locations of the 2 DF sites (X) and the limits (O) of the grid of transmitters considered for Network C. The latitude and longitude lines are given for every 10°.

As expected, these figures show that the highest productivities are obtained when two SSL sites are available and only one is required to hear the transmitter. The lowest productivities are obtained when both sites are required to hear each transmitter. Of the two sites, DF2 has the higher productivities, but this is only because of its more central location. Taking a 0 dB S/N threshold as typical for a modern DF system, during the day the productivity of two SSL systems exceeds that of two LOB systems by an average factor of 2.4. In other words, giving the systems at DF1 and DF2 a range capability as well as the usual azimuth capability allows the working of twice as many transmitters. In fact, the productivity of a single SSL system at either location well exceeds that of a 2-station LOB network. At night, the advantage is somewhat less, at 1.44. This is because HF absorption has disappeared, and the main reason for different productivities is whether or not the frequency is supported by the ionosphere on the circuit being considered. The higher frequencies would not be supported on the shorter circuits.



6. Productivities of Network C as a function of system S/N threshold, for daytime and for frequencies from 3 to 7 MHz. There are 4 cases - 1. DF1 only; 2. DF2 only; 3. DF1 OR DF2; 4. DF1 AND DF2.



7. Productivities of Network C as a function of system S/N threshold, for nighttime and for frequencies from 3 to 7 MHz. There are 4 cases - 1. DF1 only; 2. DF2 only; 3. DF1 OR DF2; 4. DF1 AND DF2.

The advantage offered by the use of SSL systems increases rapidly as the required S/N threshold increases. For example, during the day the advantage increases to 3.5 for a 5 dB threshold. During the night, the advantage increases to 2.4. In practice, all target locations would not be equally important. Examination of the variation of the productivity over the region being monitored, for a frequency range of 3 to 7 MHz, leads to the following conclusions:

#### Daytime; 3-7 MHz.

The productivities drop off with distance from the DF sites, because of the increasing effect of absorption on the longer circuits. The effect is worst for the DF1 and the DF1 AND DF2 cases, since DF1 is at one corner of the grid.

The relative advantage of the SSL approach increases to a factor of 5 for the western half of the grid, since the bulk of these transmitters can be heard only by DF2.

**Nighttime: 3-7 MHz.**

The productivities reach 100% for the DF1 OR DF2 case. HF absorption has disappeared, and the higher frequencies not supported on a short circuit to one DF site will be supported on the other.

The productivities for the western gridpoints are higher for DF2 than for DF1, because the circuit lengths to DF1 are so long that the dominant propagation mode is a 2-hop mode, with an extra ground-reflection loss.

The productivities for the western gridpoints for the DF1 AND DF2 case are low, because they cannot be heard at DF1.

## 5. SIMULATED ACCURACIES OF AN HF-DF NETWORK

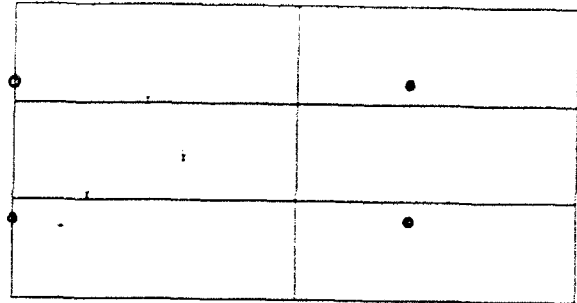
### 5.1 Introduction

This section considers by way of illustration the accuracies obtainable with HF-DF networks in several locations, with varying numbers of SSL and LOB systems. The accuracies are based on simulations, using representative bearing errors, range errors and sample sizes. The previous section has drawn attention to the importance of the concept of the productivity of an HF-DF network, which is a measure of the ability of a given DF network to work an array of transmitters spread over the area of interest, and operating on typical frequencies, with typical powers and antennas, throughout a full 24-hour period, throughout the year, and throughout a full solar cycle. These analyses have illustrated the very large advantages to be gained by the use of SSL systems, rather than just LOB systems. The source of the advantage lies in the requirement for only one SSL site to be able to hear a specified transmitter, whereas the LOB approach relies on at least two, and preferably three, sites being able to hear each transmitter.

The **accuracy** of a DF network is also an important, although secondary, consideration since position estimates with a high level of uncertainty would be counter-productive. A very high *theoretical* accuracy, based on simple geometrical considerations, is of course completely irrelevant if the required number of members of the network cannot hear the same transmitter. An analogy from physical optics which comes to mind is the intensity pattern generated from two slits - there will be an interference pattern (accuracy) generated, but only at those angles which do not correspond to minima in the diffraction pattern (productivity).

### 5.2. Method of Analysis

Three stations have been considered for Network D, as illustrated in Figure 8. As with the productivity analyses, an array of transmitters is considered, covering the pseudo-latitude range 0°N(2°)14°N and the pseudo-longitude range 0°W(2°)14°W. Observed LOBs and calculated ranges have been simulated for each transmitter and DF site by a random distribution about the correct values (known: a priori, since we know the locations of the transmitters and DF sites), with specified RMS errors and no mean error. The RMS bearing errors were allowed to vary with range as suggested by the well-known Ross curve, while the RMS range error was set at (2000/Range) percent, except that the error could not be less than 10% of range (in line with expected values).



8. Locations of the 3 DF sites (X) and the limits (O) of the grid of transmitters considered for Network D. The latitude and longitude lines are given for every 10°.

The simulated LOBs and ranges were set up by a main program which then called in a standard fixing algorithm to determine the corresponding target location. The parameters discussed in this paper are the area of the elliptical 90% confidence region, and the median percentage miss distance for the 8 x 8 transmitters (the distances between the correct locations and the Best Position Estimates (BPEs)) returned by the fixing algorithm. The effect of introducing a random error into the correct location of a transmitter (such as would be caused by site errors or tilts) is to increase the median miss distances, and to prevent them from approaching zero. There is negligible effect on the area of the ellipse. This point will not be pursued further here.

### 5.3 Results for Network D

Table 1 presents the median values of the elliptical area and miss distances for the 8 x 8 gridpoints, for different sample sizes. There are three different parameters or situations for which we need to compare the two measures of error:-

1. Sample size
2. Number of stations
3. Inclusion of range estimates.

Note that the exact values of the entries in the table are

	STATIONS	RANGES?	N=5	N=12	N=19	N=38
AREA	3	YES	256	107	67	34
	3	NO	318	132	83	42
	2	YES	497	201	127	63
	2	NO	645	272	171	84
MISS	3	YES	7.8	5.2	4.4	3.8
	3	NO	8.5	7.0	4.7	3.5
	2	YES	13.6	5.4	5.2	3.8
	2	NO	11.7	8.8	7.0	5.8

Table 1. Median values of the elliptical search areas and miss distances for different sample sizes, for Network D. The number of DF sites is either 2 or 3, and range estimates may or may not be included.

not important.

### 5.3.1 Sample Size

For an infinite sample size,  $N$ , the area of the ellipse can be expected *a priori* to approach an asymptotic limit, and this trend can be seen in the table. The decrease in area with increasing sample size is in fact quite dramatic. For example, the top line in the table shows a halving of the area for each increase in sample size. The median miss distance, on the other hand, should asymptote to zero, provided the errors introduced into the simulated observations are truly random. For example, for a 3-station fix including ranges (line 5), the median miss distance decreases from 7.8 km to 3.8 km.

The rapid increase in the area of the ellipse as the sample size decreases (read the table from right to left) highlights the fact that a small sample size is not necessarily representative of the total population of observations, and draws attention to the need for as many observations of a given transmitter to be obtained as possible.

### 5.3.2 Number of Stations

The effect of the number of stations can be determined from a fixed column in the table, taking the entries in pairs, according as range estimates have been included or not. For example, the  $N=38$  column for the first four rows shows that ignoring one station (DF3) leads to a doubling of the area of the ellipse (from 34  $\text{km}^2$  to 63  $\text{km}^2$ ; and from 42 to 84  $\text{km}^2$ ).

### 5.3.3 Inclusion of Range Estimates

The effects of including range estimates can be determined from alternate lines of the table, corresponding to YES (ranges included) or NO. The table shows that there is about a 25% increase in the elliptical area when no range estimates are available for a three-station network, and about a 33% increase for a two-station network (without DF3). As expected, the effect of not having range estimates is greater for the network with the smaller number of stations. The median value of the miss distance also increases on average when the range estimates are excluded, but the effect is small and not consistent.

The results show that the inclusion/exclusion of range estimates does not have dramatic effects, certainly not as large as those of removing a station or decreasing the sample size. This is presumably because the size of the DF net is commensurate with the size of the region being monitored, and the geometry is favorable for the LOB approach. The relative effects of the different errors also depend, of course, on the philosophy adopted in the fixing algorithm.

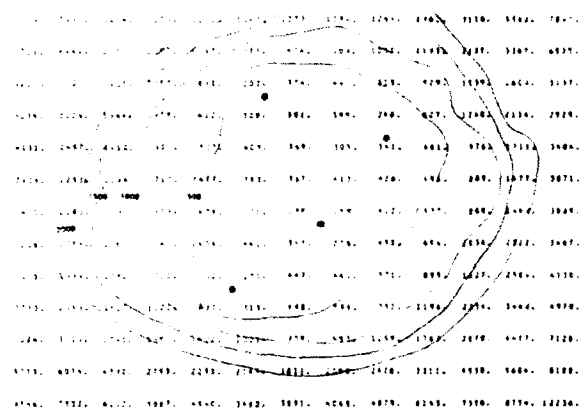
In other situations, in which the size of the DF network is much smaller than the region being monitored and the geometry becomes very unfavorable (see Section 5.6, for example), it is the accuracy of the range estimates which controls the overall accuracy of the DF network. For the present case, it is only when the geometry becomes extremely unfavorable that the use of range estimates becomes important.

The elliptical areas for each latitude/longitude gridpoint have been calculated for four cases - 3 stations, with ranges; 3 stations without ranges; 2 stations with ranges; 2 stations without ranges. The largest effects of excluding the range estimates occur for gridpoints along the DF1-DF3 axis (the south-west gridpoints); and along the DF1-

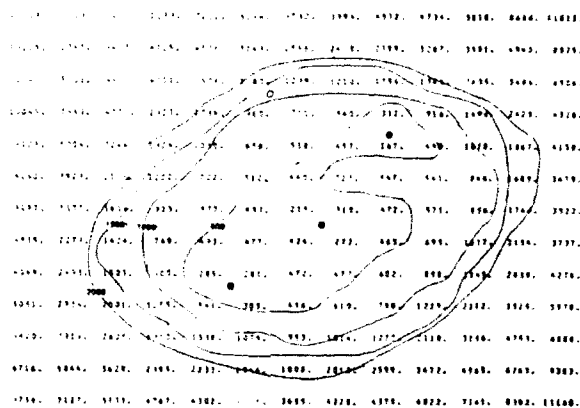
DF2 axis (the south-east gridpoints). The latter increases are particularly large when DF3 is removed, typically a factor of three.

### 5.4 Results for Network A

Grid-point maps and associated contour plots of the area of the 90% search ellipse for Network A are given in Figures 9, 10 and 11. The sample size was set at five observations of each transmitter by each station, and a daytime Ross curve was used to estimate the variance of the bearing. Figure 9 gives the results for the situation in which all four stations can hear a transmitter, Figure 10 gives the results for the three southernmost stations (excluding DF1). The contours enclose regions with search areas of 500, 1000, 1500 and 2000  $\text{km}^2$ .



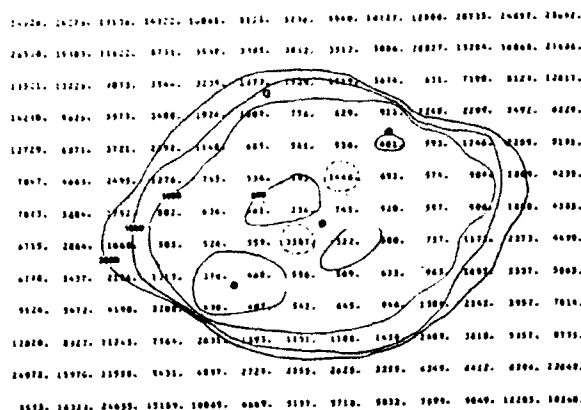
9. Calculated 90% elliptical search areas (in  $\text{km}^2$ ) for transmitters located at each of the latitude-longitude gridpoints for Network A. The gridpoints are separated by 5°. The contour lines include search areas less than 500, 1000, 1500 and 2000  $\text{km}^2$ . All four SSL sites have been included.



10. Calculated 90% elliptical search areas (in  $\text{km}^2$ ) for transmitters located at each of the latitude-longitude gridpoints for Network A. The gridpoints are separated by 5°. The contour lines include search areas less than 500, 1000, 1500 and 2000  $\text{km}^2$ . The northernmost SSL site (DF1) has been excluded.

Figure 9 presents no surprises, with 35 gridpoints having search areas less than 500  $\text{km}^2$ . When only the three southernmost stations are used (Figure 10), the region with search areas less than 500  $\text{km}^2$  decreases to 20 gridpoints. The search areas increase to over 10<sup>4</sup>  $\text{km}^2$  at the extreme edges of the area being considered.





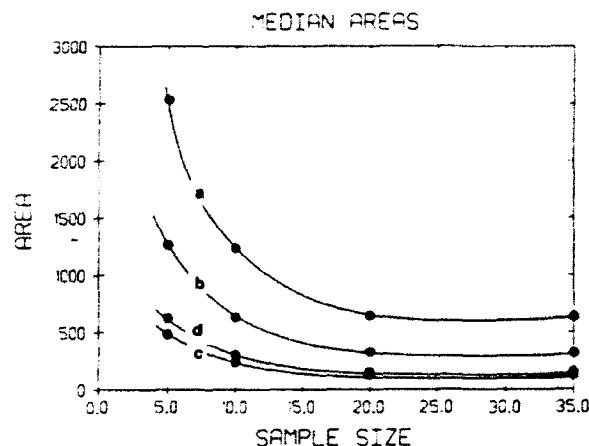
11. Calculated 90% elliptical search areas (in km<sup>2</sup>) for transmitters located at each of the latitude-longitude gridpoints for Network A. The gridpoints are separated by 5°. The contour lines include search areas less than 500, 1000, 1500 and 2000 km<sup>2</sup>. The northernmost DF site has been excluded, and range estimates have not been included.

Figure 11 presents the results for the same situation as Figure 10 (DF1 excluded), except that range estimates have not been used, only the LOBs. Without DF1, the network collapses almost to a straight line, which is a very poor choice of geometry for a DF network. As would be expected, the search areas become very large for transmitters lying along the extension of the DF baseline, to the north-east and south-west. Some very large errors arise for transmitters lying between DF2 and DF4. The errors also increase rapidly in the north-east and south-west directions for transmitters 500 km or so from DF4 and DF2. This rapid increase is exemplified by the very closely-spaced contour lines.

### 5.5 Results for Network C

There are two cases to consider here, the accuracy of each SSL site acting independently, and the accuracy of the two acting together. Following the procedures described earlier, numbers of values of azimuths and ranges have been generated, randomly distributed about the correct values, and passed to the fixing algorithm. The results are discussed here in terms of the median value of the elliptical search areas for the 8 x 10 gridpoints. Figure 12 shows how the median value decreases with sample size, for several cases - (a) DF1 only; (b) DF2 only; (c) DF1 and DF2, with range estimates; (d) DF1 and DF2, with no range estimates. Case (d) corresponds to the traditional LOB approach.

It can be seen from Figure 12 that there are two ways to reduce the size of the elliptical area, firstly by increasing the sample size, and secondly by using a second SSL site. In fact, the two procedures may be considered identical in some respects, since adding a second SSL station doubles the sample size. For a given sample size, the areas are significantly greater for DF1 than they are for DF2. This is because DF1 is at one corner of the grid, and the areas increase with the length of the circuit. When both sites can hear the transmitter, the size of the areas drops significantly. The advantage of having range estimates is not as great as being able to hear the transmitter from both sites. The size of the search area changes with the

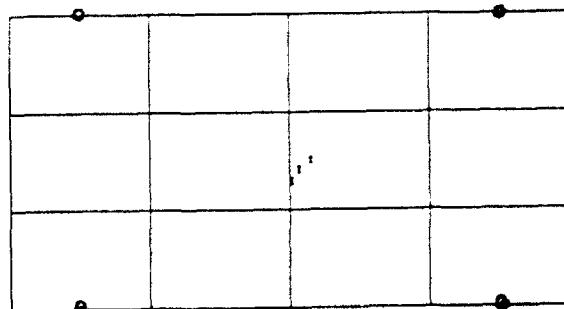


12. Median values of the elliptical search areas as a function of sample size. There are four cases - a. DF1 only; b. DF2 only; c. Both DF1 and DF2, with range estimates; d. Both DF1 and DF2, with no range estimates.

position of the gridpoint, the main effect being an increase in the area with circuit length. When range estimates are not available, the areas in the western half of the grid increase by a factor of 3. The calculations become unreliable for those gridpoints located near the extension of the DF1-DF2 baseline, because of the unfavorable geometry.

### 5.6 Results for Network E

The problem with Network E is to monitor a very large area, using several DF sites with a small baseline. The assumed locations of the DF stations are illustrated in Figure 13, along with the area to be monitored. The relatively short baseline brings with it very unfavorable geometry for a network solution based on LOBs, except for transmitters within about 500 km or so from the DF sites. The SSL approach to position estimation is therefore mandatory for most transmitters.



13. Locations of the 3 DF sites (X) and the limits (O) of the grid of transmitters considered for Network E. The latitude and longitude lines are given for every 10°.

For this particular application, the relative productivities of the LOB and SSL approaches are not important, since the S/N ratios at each DF station would be almost the same for all target transmitters because of the relatively small separation of the DF stations. This is in sharp contrast with

other scenarios for which the productivity issue is paramount. The accuracy issue is more important than the productivity issue, especially for those transmitter locations for which the available DF network offers very poor geometry for a network solution. The productivity issue would be more important for transmitters within about 600 to 1000 km of the DF sites, but these are not our main concern here.

Table 2 lists the median percentage miss distances for transmitters located at the intersections of pseudo-latitudes from 0°N to 30°N, and pseudo-longitudes from 0°E to 30°E. These results were derived using the techniques described in Section 5.2, but with one important modification. One difficulty with using simulated observations derived by randomizing the correct values is that the errors tend to average out. In an effort to overcome this effect, the program was looped through the calculations of the miss distances a number of times (typically 5), and the largest miss distance derived for each latitude/longitude gridpoint selected. A "worst case" scenario is thus built up, and this is what is given in the table.

0.0	1.5	1.8	3.0	3.8	2.3	3.2	3.3	
5.0	2.5	3.3	3.3	3.1	2.4	3.8	4.2	
0.0	1.9	2.6	1.0	3.5	2.8	3.4	3.0	3 SITES
5.0	1.6	2.8	2.5	11.6	2.3	4.3	2.0	WITH RANGES
0.0	3.2	1.2	2.3	2.4	2.3	2.1	3.4	
5.0	3.5	3.1	2.3	1.6	3.5	1.7	1.5	
0.0	7.7	3.2	1.2	1.9	3.1	3.1	2.5	
Median value of % miss distance (approx) 2.8								
0.0	25.1	21.9	18.7	41.6	47.8	86.5	99.7	
5.0	16.2	12.3	12.8	11.8	42.8	99.8	98.1	
0.0	11.1	5.4	4.6	6.4	97.5	30.5	24.8	3 SITES
5.0	9.0	4.7	6.9	11.9	6.2	5.4	10.3	WITH RANGES
0.0	14.5	11.3	12.0	11.8	3.8	8.1	7.3	NO RANGES
5.0	31.9	29.0	69.0	13.0	5.9	10.5	12.3	
0.0	80.0	73.2	82.2	24.8	15.2	18.3	19.3	
Median value of % miss distance (approx) 14.6								
0.0	1.7	6.4	1.9	1.4	5.3	3.6	7.9	
5.0	3.1	5.0	1.7	4.0	3.5	2.6	5.3	
0.0	3.1	2.0	2.3	3.0	3.4	2.3	5.2	2 SITES
5.0	2.7	4.7	1.9	10.5	3.4	3.5	4.4	WITH RANGES
0.0	2.3	2.5	1.7	2.0	3.2	2.4	3.0	
5.0	2.5	3.5	5.1	4.3	3.6	2.8	2.9	
0.0	4.0	2.7	4.6	2.3	3.7	7.4	4.5	
Median value of % miss distance (approx) 3.2								
0.0	15.9	14.2	16.7	23.9	42.9	98.4	99.9	
5.0	15.8	5.5	11.0	98.9	37.3	99.8	83.3	
0.0	14.7	7.1	2.2	6.4	73.6	13.9	21.1	
5.0	9.1	5.9	2.7	10.8	3.2	7.3	13.7	2 SITES
0.0	12.0	7.8	8.5	6.8	2.4	6.3	8.2	NO RANGES
5.0	26.6	25.5	1436.8	12.6	5.6	8.1	3.9	
0.0	51.8	79.8	77.8	20.0	11.8	15.1	8.7	
Median value of % miss distance (approx) 13.7								

Table 2. Median percentage miss distances for the 8 x 8 gridpoints of Network E. The spacing is 5°. The number of DF sites is either 2 or 3, and range estimates may or may not be included.

The table shows that the effects of including range estimates is to decrease the median miss distances for the area covered by a factor of 5. The inclusion of a third DF station (DF3) has little effect. The largest errors occur generally in the south-east and north-west of the area being monitored, with a local increase in error near the DF stations themselves. The DF stations are located approximately at the middle of each section of the table. The accuracy advantage offered by the use of range estimates reaches a factor of 10 or more in the south-west

and north-east portions of the area. The minimum errors occur for transmitters located close to the perpendicular bisector of the DF baseline, i.e., roughly from north-west to south-east.

## 6. ADVANTAGES & DISADVANTAGES OF THE LOB & SSL APPROACHES

### 6.1 Introduction

The previous sections have provided sufficient background to allow a comparison to be made between the traditional LOB technique of position estimation and the SSL technique. In fact, the two approaches are complementary.

### 6.2 Advantages of the SSL Approach

Most of the advantages of the SSL approach arise directly from the ability of an SSL site to work by itself, and the ability to work short-range targets. The advantages include

1. Only one (self-contained) station is required.
  - 1.1 This simplifies the logistics and personnel requirements, decreases the support costs, and minimizes the need for site defense and security.
  - 1.2 There is no confusion as to which transmitter is being worked.
  - 1.3 There is no need for a sophisticated system of network communications to coordinate simultaneous measurements on target transmitters.
2. With tilt correction, it is possible to work short-range transmitters.
3. The elevation angles of the incoming waves are measured and used, rather than simply ignored. The SSL procedure therefore takes greater advantage of the available information.

### 6.3 "Disadvantages" of the SSL Approach

1. For best results, an ionosonde such as the Andrew sounder TILTSONDE™ is required.

1.1 Although TILTSONDE™ is an LPI instrument, it is still an active instrument. TILTSONDE™ can thus potentially interfere with other nearby sensitive receiving equipment, and can potentially be located by an opposition SSL system. [In practice, the potential in both cases seems to be exceedingly low.]

1.2 The requirement for an ionosonde brings with it extra costs (per site), the need for extra logistic support, extra antennas and ionogram interpretation skills.

2. By itself, a SSL site cannot determine the order of the propagation mode, i.e. the number of hops. This is the hop ambiguity.

3. Even with a co-located ionosonde, it is possible for an E<sub>s</sub> propagation mode to be wrongly identified as a normal F mode, and vice versa, because of the low spatial correlation distance of E<sub>s</sub>. The presence of E<sub>s</sub> may also lead to M and N modes, which cannot be identified simply on the basis of the observed AOA. This is the E<sub>s</sub> ambiguity.

### 6.4 Advantages of the LOB Approach

1. There are no hop or E<sub>s</sub> ambiguities.
2. Especially for long-range transmitters, when the uncertainty in range estimates can become quite large compared to the effects of a typical bearing error, the network approach is able to provide more accurate position estimates than the SSL approach. [However this superior accuracy is predicated on a reasonable geometry and on more than one DF site being able to hear the transmitter.]

## 6.5 Disadvantages of the LOB Approach

1. Increased costs. Although each site may be cheaper than an SSL site, more sites are required. This brings with it all of the features arising from having to maintain more than one site - assuming that conditions are such that more than one suitable site is available in the first place.

2. The need for network communications. An LOB site cannot work alone, and all results rely on the exchange of information around the network. This information will include such things as the frequency and modulation of the target transmitter, any identifying information, and the LOBs observed from each site which can hear the transmitter.

3. By themselves, LOBs are completely useless for short-range targets, since they are so heavily affected by the local ionospheric tilts.

4. The productivity of a network can be very low, especially when 3 or 4-site fixes are required.

5. The approach needs favorable geometries, preferably with the LOBs cutting at right angles. When all available LOBs are similar, such as when the available baseline is short compared to the area being monitored, very large errors are likely [unless estimates of the ranges are also available].

## 6.6 Some Remedies

Provided a reasonable geometry is available, the main problem with the LOB approach is that of productivity. Put simply, the LOB approach requires that the frequency of the target transmitter lies between the LUF and MUF for the circuits to all DF sites. If the frequency is too high on the shorter circuits, nothing can be done, since the DF site lies inside the skip zone. However if the problem is one of the target's frequency suffering from high amounts of absorption, the possibility exists for the signals to be processed using modern DSP techniques.

The disadvantages of the SSL approach include the use of an active sounding instrument (ionosonde), and the E, and hop ambiguities if only one SSL site is available. If the former is found to be a problem, the possibility exists for extensive monitoring of check targets [11], in order to set up an effective VI ionogram, or to update an ionospheric index or effective sunspot number. This approach can tie up the AOA assets, but it would provide a more accurate position estimate than the use of synoptic models which are not updated on the basis of local observations.

The E, and hop ambiguities can often be removed by the use of ancillary information, such as language, and national and natural boundaries. In general, position estimates for an observed elevation angle should always be given in terms of (1) an E, propagation mode (including at least N modes), (2) a normal propagation mode, and (3) single and double-hop propagation modes.

## 7. CONCLUSION - THE OPTIMUM SOLUTION

Generally speaking, the SSL approach is the only one possible for short-range targets, or when only one DF site can hear the transmitter; the SSL and LOB approaches should be about equally accurate for medium-range targets; and the LOB approach should be more accurate for long-range targets, especially for multi-hop circuits. However in this last case, the low productivities become an important issue.

It does not take much thought to recognize that the optimum solution is to have all existing HF-DF networks upgraded to HF-SSL networks. Any new networks (of one or more sites) should be made SSL networks automatically, with the appropriate number and locations of the nodes being determined by simulation, in conjunction with extraneous constraints. The LOB and SSL approaches would then nicely complement each other.

A network of SSL sites can not only perform all the standard functions of an LOB network, but also offers the advantages of being able to work short-range transmitters, of offering higher productivities, and of being able to provide a solution when only one DF site can hear a transmitter. Networks of SSL sites are also more stable against loss of a member than are networks of LOB sites. In fact, one SSL site can in many situations work more targets than two or three LOB sites.

## 8. REFERENCES

1. McNamara, L.F. *The Ionosphere : Communications, Surveillance and Direction Finding*. Krieger Publishing Co, Melbourne, Florida, 1991.
2. McNamara, L.F. *Determining the Location of an HF Transmitter from Observations at a Single Site of Angles of Arrival of Skywaves from that Transmitter*. Andrew SciComm Internal Report, January 1991.
3. Davies, K. *Ionospheric Radio*. Peter Peregrinus Press, London, 1990.
4. Gething, P.J. *Radio Direction Finding*. Peter Peregrinus Press, on behalf of the IEEE, London, 1978.
5. Wilkinson, P.J. *Monthly Ionospheric Indices*. Proceedings of the Workshop on Solar Terrestrial Physics, Meudon, June 1984, pp 505-508.
6. Chen, J., J.A. Bennett and P.L. Dyson. *Automatic Fitting of Quasi-Parabolic Segments to Ionospheric Profiles with Application to Ground Range Estimation for Single-Station Location*. J Atmos Terr Phys, 52(4), pp 277-288, 1990.
7. Jones, R.M. and J.J. Stephenson. *A Versatile Three-Dimensional Ray Tracing Computer Program for radio Waves in the Ionosphere*. Office of telecommunications Report OT-75-76, Department of Commerce, Washington, 1975.
8. McNamara, L.F. *Ionospheric Modelling in Support of Single Station Location of Long-Range Transmitters*. J Atmos Terr Phys, 50(9), pp 781-795, 1988.
9. McNamara, L.F. *Analysis of Angle of Arrival Observations of WWV (Fort Collins) Transmissions Received in Dallas, Texas*. Andrew SciComm Internal Report, November 1991.
10. Teters, L.F., J.L. Lloyd, G.W. Haydon and D.L. Lucas. *Estimating the Performance of Telecommunication Systems Using the Ionospheric Transmission Channel*. NTIA Report 83-127, U.S. Department of Commerce, July 1983.
11. McNamara, L.F. *Real-Time Ionospheric Specification Using Passive Monitoring of Angles of Arrival of Signals from Uncooperative HF Transmitters*. Andrew SciComm Internal Report, January 1991.

## DISCUSSION

## P. GEORGE

I have a difficulty with your table of median values of search areas and distances for network D. I agree that operational experience shows that SSL at each site increases the productivity. However, operational experience of more than 30 years also shows that inclusion of range estimates degrades the line of bearing position estimates by introducing greater variance. My question is "what range accuracy have you assumed in calculating the table?" I agree with respect to productivity but have serious trouble with accuracy. You need to look at this table.

## AUTHOR'S REPLY

I agree that the inclusion of range estimates will degrade a fix for long-range targets. However, for the simulations, I used range errors which did not exceed 10% of range. For medium-range circuits, for which a 10% range error is reasonable, an SSL fix is just as accurate as a two-station LOB cut. The ellipses in this case are nearly circular.

## C. GOUTELARD

La localisation par mesure d'azimut présente un avantage sur la localisation SSL en guerre électronique. Il est en effet plus difficile pour un émetteur ennemi de brouiller simultanément plusieurs sites de réception. Il rencontre en effet la même difficulté que vous en ce qui concerne la propagation ionosphérique. Ne pensez-vous donc pas que ceci soit un argument supplémentaire pour conserver les deux systèmes dans les projets stratégiques?

*In electronic warfare, azimuth measurement has an advantage over SSL localization. It is in fact more difficult for an enemy emitter to simultaneously jam several receiving sites. Actually, the enemy emitter encounters the same problems with ionosphere propagation as you do. Don't you think that this could be an additional point in favor of keeping both systems in strategic projects?*

## AUTHOR'S REPLY

You are quite correct. It is much harder for a transmitting site to avoid being intercepted by two sites than by one site, so two DF sites are clearly better than one. A judicious choice of operating frequency can in fact make things hard for a single SSL system to provide an accurate position estimate.

## LEROUX

Comment, pour un site faible, discriminez-vous une onde de sol d'une onde ionosphérique très éloignée?

*For a low elevation angle, can you discriminate a ground wave from a sky wave? In one case  $D < 70$  km, in the other  $D \geq 1000$  km.*

## AUTHOR'S REPLY

A skywave and groundwave can often be differentiated by an operator because of the fading which exists on the skywave. It may also be possible to distinguish between them using the quality factors ascribed to each angle of arrival measurement.

## R. ROSE

I agree with your general comments about SSL and their shortcomings. However, on a global basis, my experience with BULLSEYE shows that you need as many sites as possible. In limited areas, such as designing a net to cover Australia, netted SSL may be appropriate. On a global basis, netted azimuth-only systems with large wide aperture antennas are more appropriate.

## AUTHOR'S REPLY

The point I make in my paper is that large wide-aperture systems are becoming less available in the current era. I also find it hard to believe that the world-wide nets do not suffer from low productivities.

## HIGH RESOLUTION SAMPLED APERTURE ARRAY DIRECTION FINDING AT HF

T.N.R. Coyne

Defence Research Establishment Ottawa  
Ottawa, Ontario, K1A 0Z4  
CANADA

## SUMMARY

A simple parametric target model fitting algorithm, to give a least squares fit to the observed in-phase and quadrature data across a linear sampled aperture array, is outlined. The performance of this algorithm with various simulated signals and noise levels is presented. In particular, it is shown that this approach is quite robust in the presence of noise. The algorithm was developed for the analysis of data obtained on a large HF sampled aperture receiving array operated near Ottawa. An example of this analysis is given. In this instance a long range signal exhibiting typical ionospheric fading (on individual array elements) is shown to be resolvable into three stable, coherent components spread over a  $2.5^\circ$  arc.

## LIST OF PRINCIPAL SYMBOLS

s	suffix denoting array element, $-w \leq s \leq w$
t	suffix denoting incident wave, $1 \leq t \leq v$
$\lambda$	wavelength
d	element spacing
D	element spacing in radians, i.e. $2\pi d/\lambda$
$a_t$	incident wave amplitude
$\beta_t$	incident wave phase, at centre element
$\alpha_t$	incident wave direction of arrival, measured from boresite
$P_{s,t}, Q_{s,t}$	in-phase and quadrature components of incident wave t at element s
$P_s, Q_s$	observed in-phase and quadrature components at element s
$A_s, B_s$	observed amplitude and phase at element s
(.)	initial point for Taylor's expansion

## 1. INTRODUCTION

Between 1969 and 1980 the Communications Research Centre, sponsored by the Canadian Department of National Defence, operated an extensive High Frequency Direction Finding (HFDF) research programme. The objective of this programme was to investigate and evaluate the effects of the ionosphere on DF accuracy. To this end, an advanced

Sampled Aperture Receiving Array (SARA) was developed near Ottawa. This was used in conjunction with a portable beacon transmitter operated at various sites in Canada and the USA to study bearing errors on a wide range of propagation paths. A detailed description of this instrumentation has been given by Rice and Winacott (Ref. 1).

The principal mode of experimental operation and analysis was to transmit a signal that would allow the separation of multiple ionospheric paths (if present). Measurement of the direction of arrival of individual paths could then be made by a spatial Fourier analysis, or more commonly, by a least squares straight line fitting to the observed phase variation across the array (Ref. 2). Path separation was achieved by exploiting either range or Doppler variations. In the former case by using a swept frequency CW signal, and in the latter, by a CW signal with dwell times long enough to permit resolution.

A variety of algorithms giving angular (or in a more general sense spectral) resolution superior to that achieved by conventional Fourier-based analysis, have appeared in recent years. Perhaps the best known of these is the MUSIC algorithm (Ref. 3). Several comparative performance studies and reviews of these techniques have been published (see, for example, Ref. 4 and Ref. 5). Of the various algorithms the Parametric Target Model Fitting (PTMF) approach appears well suited to HFDF requirements because of its ability to handle correlated targets, a requirement if ionospheric path separation is to be achieved on general signals when recourse to other methods, such as mentioned previously, cannot be made.

During the experimental lifetime of the Canadian HFDF programme, very little effort was directed toward the application and evaluation of these "superresolution" techniques, as they were only attracting serious attention near the end of the programme. Since that time, however, the development and testing of an analysis algorithm based on the PTMF approach has been carried out. This method, referred to here as the Least Squares Fit (LSF) algorithm, gives the set of incident plane waves that provides a match, in a least squares sense, to a (single) set of in-phase and quadrature data across a sampled aperture array. Such a "snap-shot" will be referred to as a frame of data. This type of analysis is well suited to the important problem of direction finding on very short duration signals.

## 2. LEAST SQUARES FIT ALGORITHM

Consider the linear array shown in Figure 1. The in-phase and quadrature components of incident wave  $t$  at element  $s$  are:

$$p_{s,t} = a_t \sin(\beta_t + sD \sin \alpha_t)$$

$$q_{s,t} = a_t \cos(\beta_t + sD \sin \alpha_t)$$

In general, there is no solution for the  $3v$  wave variables that will provide a match to the  $2(2w+1)$  observed phase and quadrature components since:

$$2(2w+1) > 3v$$

Therefore a solution is sort that provides a match in a least squares sense, i.e., one that minimizes the error:

$$ERR = \sum_{s=-w}^w \left[ \left( p_s - \sum_{t=1}^v p_{s,t} \right)^2 + \left( q_s - \sum_{t=1}^v q_{s,t} \right)^2 \right] \quad (1)$$

This expression is not a simple second order function of the wave variables and a direct least squares analysis cannot be performed. An iterative procedure, in which a starting point in the wave variable space is chosen and progressively moved to the point of minimum error, must be employed. For the exploratory work described in this paper, where performance not computational efficiency is at issue, the procedure outlined below has been utilized.

At a chosen starting point, Taylor's expansion is used to form a limited linear relationship between the wave variables and their phase and quadrature components.

$$p_{s,t} = \left( \frac{\partial p_{s,t}}{\partial a_t} \right)_0 \delta a_t + \left( \frac{\partial p_{s,t}}{\partial \beta_t} \right)_0 \delta \beta_t + \left( \frac{\partial p_{s,t}}{\partial \alpha_t} \right)_0 \delta \alpha_t + (p_{s,t})_0 \quad (2)$$

$$q_{s,t} = \left( \frac{\partial q_{s,t}}{\partial a_t} \right)_0 \delta a_t + \left( \frac{\partial q_{s,t}}{\partial \beta_t} \right)_0 \delta \beta_t + \left( \frac{\partial q_{s,t}}{\partial \alpha_t} \right)_0 \delta \alpha_t + (q_{s,t})_0 \quad (3)$$

$$\begin{aligned} \text{where } a_t &= (a_t)_0 + \delta a_t \\ \beta_t &= (\beta_t)_0 + \delta \beta_t \\ \alpha_t &= (\alpha_t)_0 + \delta \alpha_t \end{aligned}$$

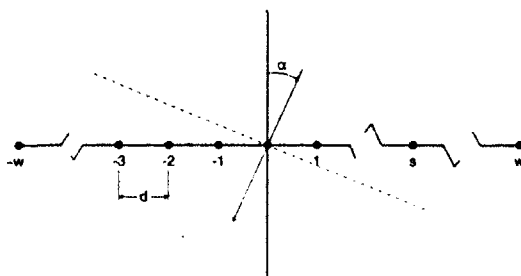


Fig. 1. Linear array of  $(2w+1)$  elements

Substituting (2) and (3) into (1) gives a second order error function that, around the starting point, approximates (1). Doing a standard least squares analysis (i.e. equating the partial differentials with respect to each of the delta wave variables to zero) on this approximate error function, yields the set of equations

$$M'M x = M'v \quad (4)$$

where

$$M = \begin{bmatrix} \left( \frac{\partial p_{w,1}}{\partial a_1} \right)_0 & \left( \frac{\partial p_{w,1}}{\partial \beta_1} \right)_0 & \left( \frac{\partial p_{w,1}}{\partial \alpha_1} \right)_0 & \left( \frac{\partial p_{w,v}}{\partial a_v} \right)_0 \\ \left( \frac{\partial q_{w,1}}{\partial a_1} \right)_0 & & & \\ \left( \frac{\partial p_{w,1}}{\partial a_1} \right)_0 & & & \left( \frac{\partial q_{w,v}}{\partial \beta_v} \right)_0 \end{bmatrix}$$

$$x = \begin{bmatrix} \delta a_1 \\ \delta \beta_1 \\ \delta \alpha_1 \\ \delta a_v \end{bmatrix} \quad v = \begin{bmatrix} p_w - \sum_{t=1}^v (p_{w,t})_0 \\ q_w - \sum_{t=1}^v (q_{w,t})_0 \\ q_w - \sum_{t=1}^v (q_{w,t})_0 \end{bmatrix}$$

The Gauss-Seidel iterative procedure is then applied. This procedure is convergent for simultaneous equations having a positive definite matrix form, such as (4), and, most importantly, the parent error function decreases monotonically with each iteration (see, for example, Ref. 6). The iterative procedure is allowed to continue until the current iterative solution reaches a boundary set by the accuracy of the Taylor's expansions. These expansions and equation (4) are then re-evaluated at this point and the iterative procedure is restarted. This continues until a point is reached where the solution varies less than some prescribed amount over a given number of iterations, the required error minimum is then considered to have been reached.

In common with many problems of this type, since the error function (1) is not a simple second order function of the wave variables it does not necessarily have a single minimum. In general, a true minimum and a number of pseudo minima will exist. It is, therefore, necessary that the starting point be chosen such that the iterative solution path will not encounter a pseudo minimum or it will hang up there.

To establish the starting point, a conventional summed beam is formed from the in-phase and quadrature data and swept over the directions of interest. The amplitude and phase of the summed beam, phased to point in the direction  $\alpha$ , are given by:

$$AMP(\alpha) = \sqrt{\left(\sum_{s=-W}^W A_s \sin(B_s + \gamma)\right)^2 + \left(\sum_{s=-W}^W A_s \cos(B_s + \gamma)\right)^2} \quad (5)$$

$$PHASE(\alpha) = \tan^{-1} \left[ \frac{\sum_{s=-W}^W A_s \sin(B_s + \gamma)}{\sum_{s=-W}^W A_s \cos(B_s + \gamma)} \right] \quad (6)$$

$$\begin{aligned} \text{where } A_s &= \sqrt{P_s^2 + Q_s^2} \\ B_s &= \tan^{-1}(P_s/Q_s) \\ \gamma &= -sD\sin(\alpha) \end{aligned}$$

In practice a raised cosine weighting is incorporated in (5) and (6), this has been omitted here for clarity. Amplitude, phase and direction of arrival starting points are taken at directions where the swept beam exhibits clear amplitude peaks.

As well as the starting point(s), the number of waves,  $v$ , in the solution must be determined. A lower limit to the value of  $v$  is the number of clear peaks determined by the swept beam algorithm. However, if incident waves are sufficiently close in their direction of arrival as to be unresolved by the swept beam,  $v$  may be larger than this. It is then necessary to progressively increase  $v$  above the lower limit until the error function minimum does not significantly decrease further. This will be illustrated in the next section.

### 3. LSF ALGORITHM PERFORMANCE

To demonstrate the characteristics and performance of the LSF algorithm, a simple linear array with 11 elements at a half-wave spacing, is considered. This array has an unweighted half-power beam width of  $9.5^\circ$  which increases to  $12^\circ$  when a raised cosine weighting is applied. These beam patterns are shown in Figure 2. The resolution capability of this array under the conventional swept beam analysis, (5) and (6), is a function of both the relative amplitudes and phases of the incident

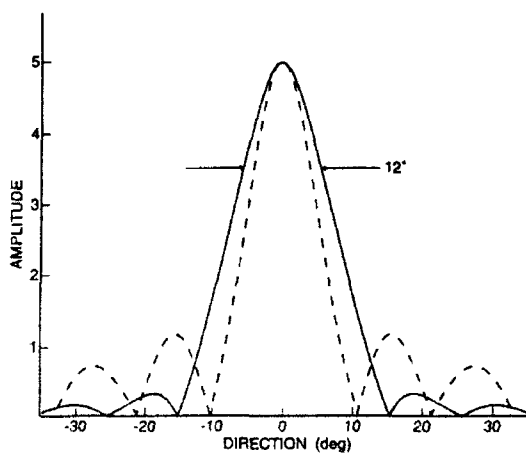


Fig. 2. Unweighted (broken line) and weighted (solid line) beam patterns of a linear array of eleven, half-wave spaced elements.

signals. This is illustrated in Figure 3 for three pairs of incident waves of equal amplitude,  $12^\circ$  directional separation and  $0^\circ$ ,  $90^\circ$ , and  $180^\circ$  phase differences. At this angular separation the  $0^\circ$  waves are unresolved, the  $90^\circ$  waves are just distinguishable as separate peaks and the  $180^\circ$  waves are clearly resolved.

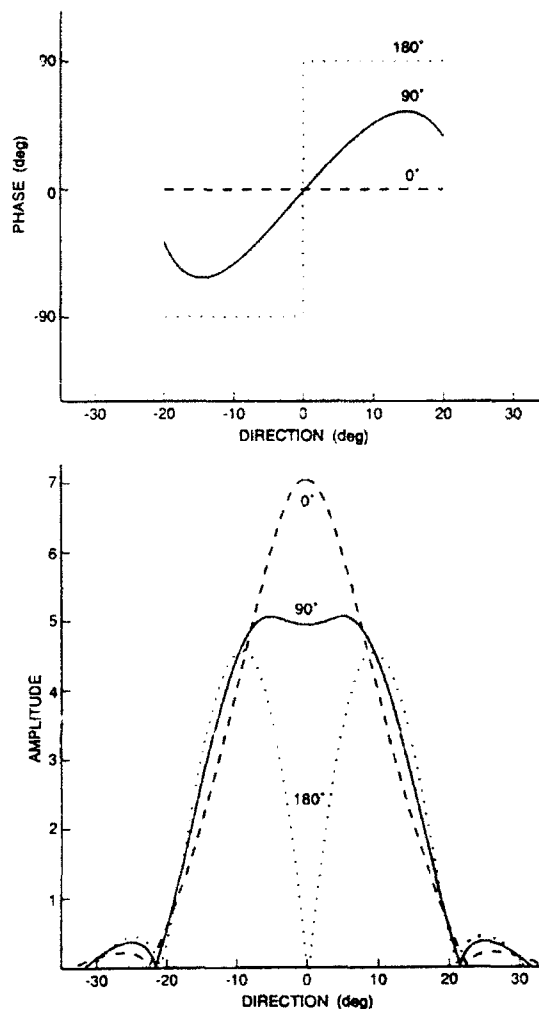


Fig. 3. Amplitude (lower) and phase (upper) of the summed beam as a function of direction for two incident waves at  $+6^\circ$  and  $-6^\circ$ , with equal amplitudes and phase differences as shown.

The question of the range of starting points for the LSF algorithm is not easily answered because the error, being a function of the  $3v$  wave variables, is not easily studied. The behaviour of the error function for the simplest case, a single incident wave is shown in Figure 4. Here the root normalized error, i.e.:

$$\sqrt{ERR/2(2W+1)}$$

is plotted as a function of two of the three variables (the third being set to its correct value). It can be seen that over a range greatly in excess of the resolution of the swept beam analysis, the error

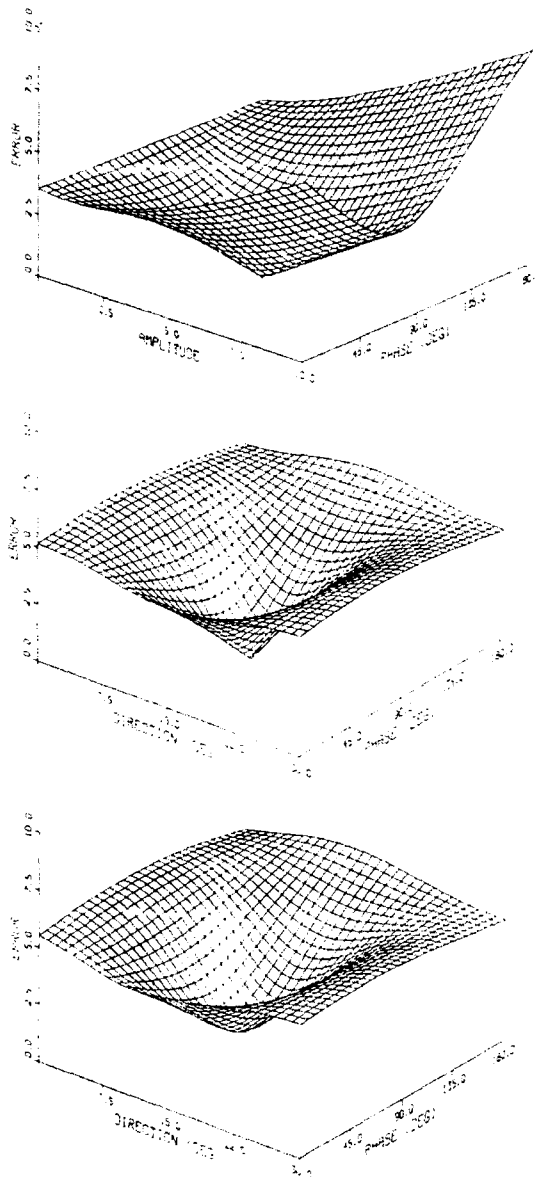


Fig. 4. Root normalized error as a function of amplitude and phase (upper) and direction and phase (middle and lower) for a single incident wave with:-

$$\begin{aligned} a &= 5.0 \\ \beta &= 75.0^\circ \\ \alpha &= 15.0^\circ \end{aligned}$$

In each case the third variable has been set to its correct value. In the lower figure noise has been added to the in-phase and quadrature data to give a 9.34 dB signal-to-noise ratio.

function is well behaved (i.e. does not exhibit pseudo minima) and the proper convergence of the iterative procedure is assured.

For two or more incident waves the error is a complex function of many variables and its study is beyond the scope of this paper. It can be shown however, that, whether or not the waves are resolved by the swept beam analysis, the starting points taken from it are well within the range of variable space where the error function is well behaved.

Like most deconvolution techniques the true value of the algorithm lies in its performance in the presence of noise. A variety of evaluations involving the simulating of both internal (receiver) and external noise have been conducted. It is outside the scope of this paper to document these, however the simple simulations shown in Figures 5 and 6 and Table 1 are judged to be very typical. Here pairs of normally distributed random numbers of the form:

$$R1 = K \cos(2\pi J) \sqrt{-2 \log_e(I1)}$$

$$R2 = K \sin(2\pi J) \sqrt{-2 \log_e(I1)}$$

where

I, J = uniform random numbers between 0 & 1

K = standard deviation

are added to each of the eleven in-phase and quadrature pairs and the resultant frame of data analyzed. Fifteen such analyses have been carried out for each of a number of single and double incident wave scenarios and nominal noise levels. (Actual noise levels produced by the random number generator are shown in the figures.) The effect of noise on the deduced phase and direction of a single incident wave is shown in Figure 5. It can be seen that on an individual frame basis, an accuracy of about one tenth of a beam width can be achieved at a 10 dB signal to noise ratio. If averaging across frames is introduced then much higher accuracy is achieved.

Two incident waves with directional separations of 5° and 2° are shown in Figure 6. These waves have a 90° phase difference and would be unresolved with conventional beam forming (compare Figure 3). Comparison of Figures 5 (lower) and 6 (upper and middle), which all have the same noise level, shows a progressive decrease in the accuracy of the derived signals. This is attributable to the interaction of the waves. As the incident waves come closer, and become less well resolved, (4) becomes more ill-conditioned and so more susceptible to noise. To achieve a tenth of a beam width resolution (on a single frame basis) about a 30 dB signal to noise ratio is required.

As incident waves can be unresolved under the swept beam analysis used to determine the starting point, the number of waves that should be used in the LSF algorithm solution is uncertain. The correct number of waves can be determined from the



residual error, i.e. the minimum error observed at the solution point. The residual root normalized error for the various simulations described above and some additional ones for waves with a directional separation of  $1^\circ$  are shown in Table 1.

The first line of the of the table is the rms in-phase and quadrature noise which is equivalent to the residual error when no waves are present (i.e. background noise). The values in brackets are the residual errors for single wave solutions to the two wave simulations.

It can be seen that when the correct number of waves is used in the solution the residual error is about at the back ground noise level (actually slightly lower as the solution to some extent "accommodates" the noise). When too few waves are used in the solution, the residual error is significantly higher. If the back ground noise is sufficiently high as to mask this, then in fact the noise is too high to allow the waves to be resolved. For example, in Table 1 the waves with  $1^\circ$  separation could not be resolved for the 9.34 or 19.34 dB signal-to-noise ratio cases.

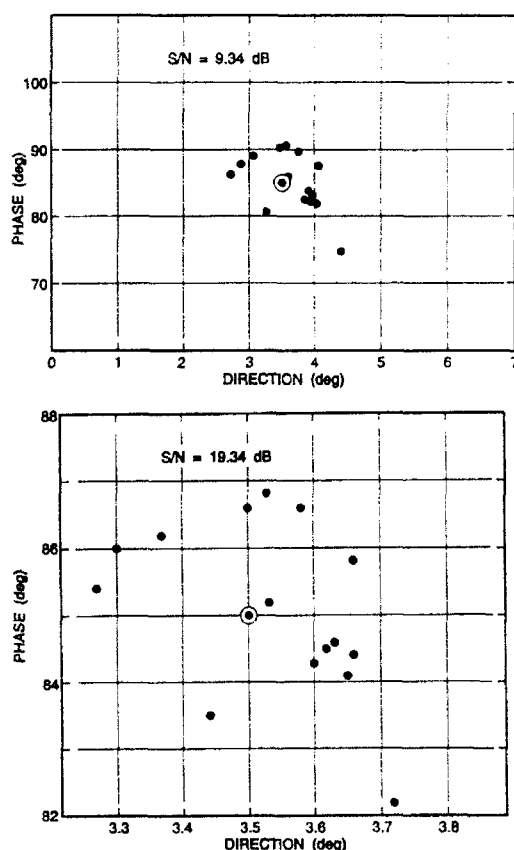


Fig. 5. Variations in derived phase and direction of arrival for a single wave with a signal-to-noise ratio of 9.34 dB (upper), and 19.34 dB (lower).

The solution has too many waves when the addition of another wave does not significantly reduce the residual error. The solution in this case is generally not erroneous, one or more of the waves simply splits into two components with the same direction of arrival.

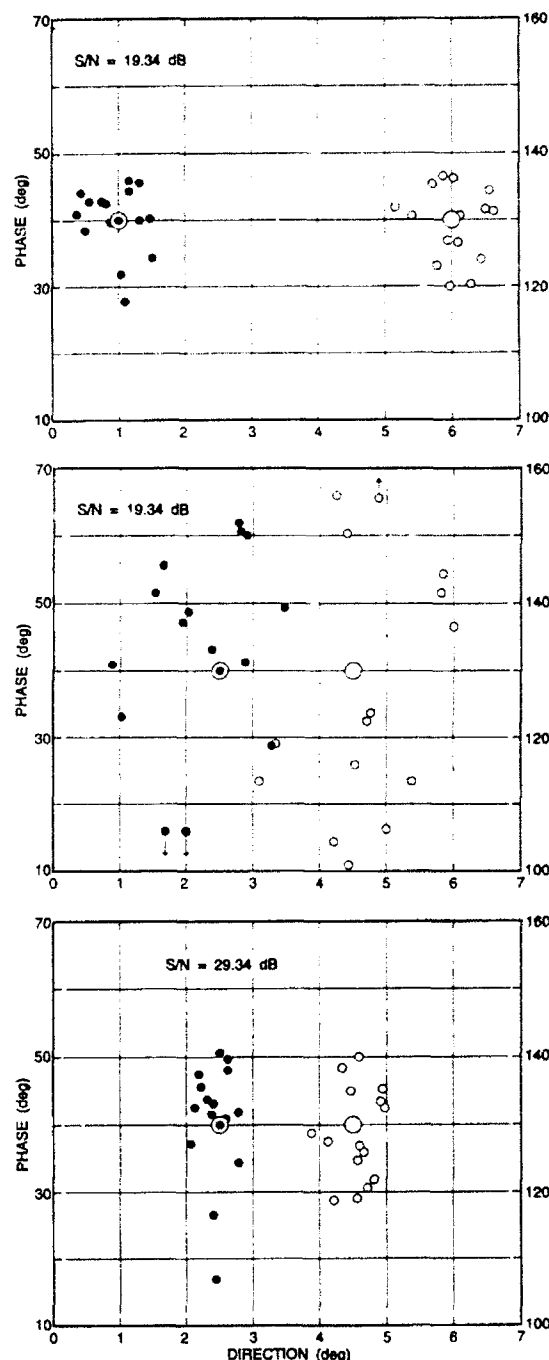


Fig. 6. Variations in derived phase and direction of arrival for two waves of equal amplitude and a  $90^\circ$  phase difference. Directional separations are  $5^\circ$  (upper) and  $2^\circ$  (middle and lower). Signal-to-noise ratios are as shown.

SIGNAL TO NOISE ----- WAVES	9.34 dB	19.34 dB	29.34 dB	39.34 dB	$\infty$
	1.206	0.381	0.121	0.038	0
3.5°	1.130	0.358			0
1° \ 6°		0.328 (2.095)			0 (2.084)
2.5° \ 4.5°	1.042 (1.380)	0.329 (0.911)	0.104 (0.860)		0 (0.860)
3° \ 4°	1.026 (1.188)	0.330 (0.544)	0.105 (0.440)	0.035 (0.431)	0 (0.4332)

TABLE 1. Residual root normalized error for various simulations.

#### 4. APPLICATION TO HFDF

The SARA facility, developed in the 1968 to 1972 time frame by the Communications Research Centre, lies some 15 km SW of Ottawa on very flat unobstructed terrain. The array consists of two orthogonal arms. The long arm has 58 antenna element positions (foundation and cabling), a central portion of 32 positions at a spacing of 7.62 meters (25 ft), one position at 13.81 meters (50 ft) off each end of the central portion, and 12 positions at 38.10 meters (125 ft) outside these. The short arm has 32 positions similar to the central portion of the long arm. For the reception of obliquely-propagated signals elevated feed monopoles are used, however these may be replaced by loops for the reception of near vertical incident signals. To minimise distortion of the incident wave field all the instrumentation and support facilities are housed in an underground laboratory located beneath the ground screen near the junction of the two arms of the array. A complete description of the facility is contained in Ref. 1.

For detailed analysis in this paper a short section of 21.88 Mhz CW data was selected. This was recorded during cooperative experiments with the Naval Ocean Systems Centre (NOSC), when the beacon transmitter was located near San Antonio, Texas, giving a range of 2350 miles and a bearing of 6.70° off the long array bore site. For these experiments only half the central antenna positions were populated, giving the 58 element lay out shown in Figure 7. At 21.88 MHz the spacing of the closest elements (central section) is just over one wavelength and spurious side lobes (or directional ambiguities) exist. However the closest of these are approximately 21° off the main beam and are not a problem when data from a known source is under study. A more positive aspect of this large element spacing, is that mutual coupling, ignored on this study, will be significantly reduced.

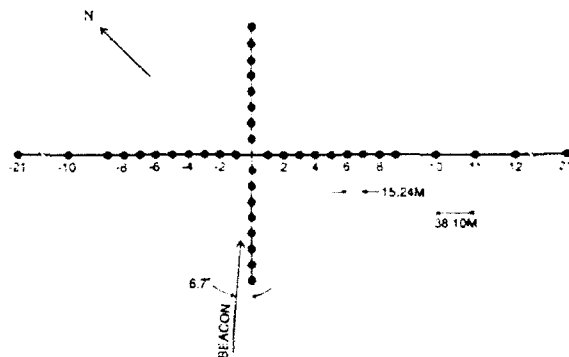


Fig. 7. Element layout of SARA facility.

Frames of in-phase and quadrature data from the 58 receivers were recorded approximately every 78 msec and the analysis of some 225 consecutive frames over about 18 sec is described here. The signal amplitude recorded on three individual elements of the array is shown in Figure 8. These show typical fading patterns which are delayed relative to each other in the manner of ionospheric "drift" data.

Application of the LSF algorithm, to data from the long arm of the array, shows the incident signal can be clearly resolved into three components with amplitude, phase variation and direction of arrival as shown on Figure 9. In performing this analysis a measure of automation was introduced, in so far as the solution for one frame of data was used as the starting point for the next. Only on the first frame of data was the swept beam analysis used. It can be seen that the three components or ionospheric paths are quite stable. The directions of arrival show little fluctuation, the cumulative phase variations are approximately linear (corresponding to constant Doppler shifts), and the amplitudes show nothing of the deep interference fades that occur when the components are not separated as in Figure 8.

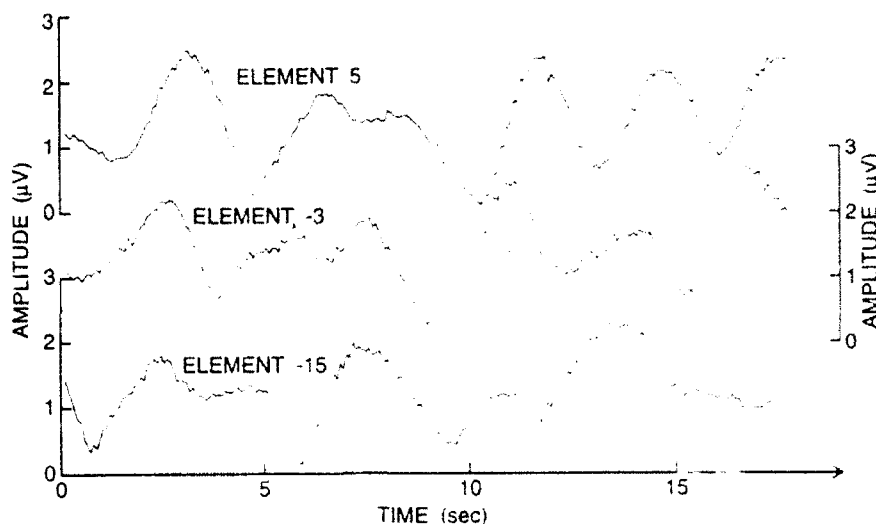


Fig. 8. Amplitude variation observed on individual elements. The variation on element 5 is delayed roughly 1 sec relative to element -3 and 4 sec relative to element -15.

The analysis of this data by wavefront testing, when the components are first separated by an FFT analysis of the Doppler shifts, has also been carried out. The analysis was over 20 sec (256 frames) giving a Doppler resolution of 0.05 Hz. Three large, and a number of smaller Doppler peaks, were observed. The characteristics of the three principal peaks, which showed good phase front linearity, are given in Table 2. The agreement with the LSF algorithm is quite good.

DOPPLER PEAK	DOPPLER SHIFT (Hz)	AMPLITUDE (uV)	DIRECTION (Deg)
1	-0.05	1.20	-7.1
2	0.15	0.47	-5.5
3	0.30	0.43	-4.5

TABLE 2. Amplitude and direction of arrival of the three principal Doppler peaks.

The LSF algorithm may be regarded as a deconvolution procedure, providing a direction of arrival resolution superior to that of the basic instrument, as represented by the swept beam analysis. A good visual evaluation of the LSF algorithm performance is obtained by comparing the swept beam analysis on the observed in-phase and quadrature data to that on the in-phase and quadrature components calculated from the LSF algorithm solution. This is shown for data frames 31 and 136 in Figure 10. The weighted beam width is about  $0.95^\circ$ , so that while component 1 is always clearly resolved under the swept beam analysis, components 2 and 3 are generally not (as in frame 31). They are, however, clearly resolved under the LSF algorithm.

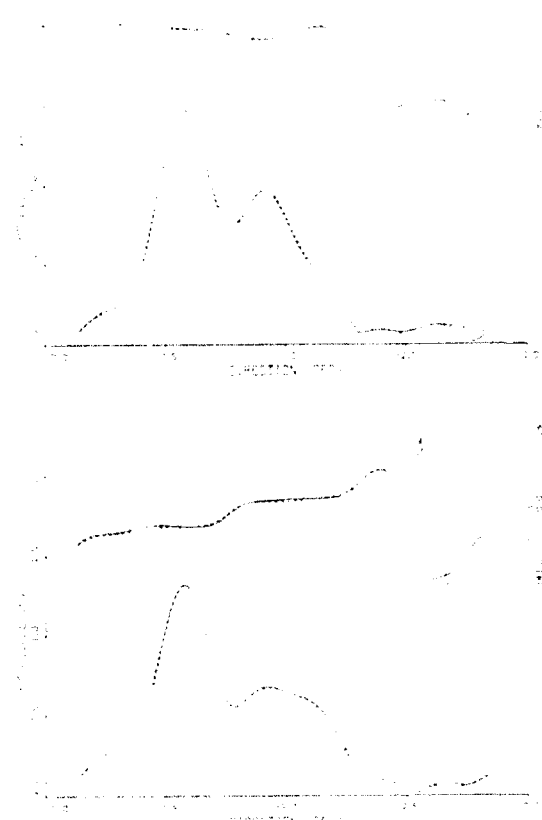


Fig. 10. Summed beam amplitude and phase as a function of direction from the observed in-phase and quadrature data (solid lines) and that derived from the LSF algorithm solutions (broken lines), for frames 31 (upper) and 136 (lower).

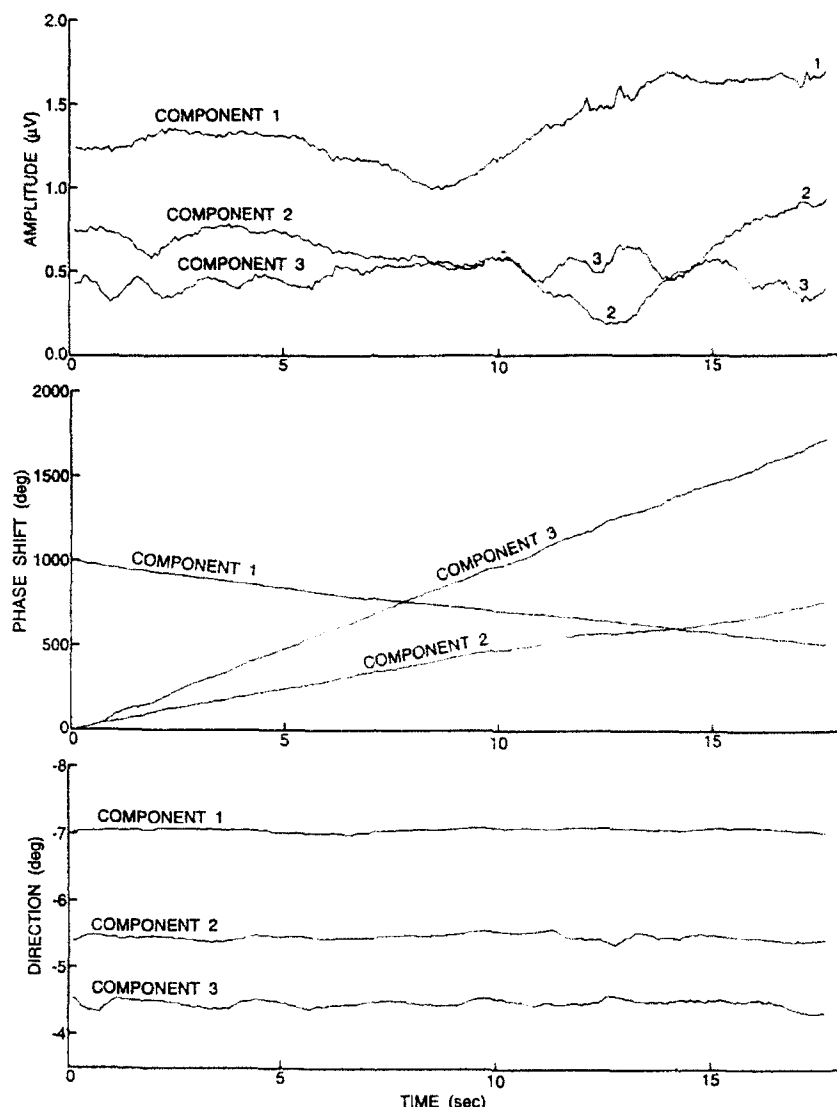


Fig. 9. Variation in amplitude (upper), phase (middle) and direction of arrival (lower) of the three components of the incident signal.

#### ACKNOWLEDGEMENTS

Mr. B.W. Lisson conducted the SARA experimental work. Mr. E.L. Winacott carried out the basic data reduction and Doppler analysis. Their assistance is gratefully acknowledged.

#### REFERENCES

1. Rice, D.W. and E.L. Winacott, "A Sampling Array for HF Direction Finding Research", Communications Research Centre Report No. 1310, Ottawa, Canada, November 1977.
2. Rice, D.W., "HF Direction Finding by Wavefront Testing", Communications Research Centre Report No. 1333, Ottawa, Canada, March 1980.
3. Schmidt, R.O., "Multiple Emitter Location and Signal Parameter Estimation", IEEE Trans., AP-34, pp. 278-280, 1986 (reprint of RADC Spectrum Estimation Workshop, October 1979).
4. Nickel, U., "Angular Superresolution with Phased Array Radar: a Review of Algorithms and Operational Constraints", IEE Proceedings, Vol. 134, Pt.F, No.1, pp. 53-59, 1987.
5. Barabell, A.J., J. Capon, D.F. DeLong, K.D. Senne and J.R. Johnson, "Performance Comparison of Superresolution Array Processing Algorithms", Project Report TST-72, Lincoln Laboratory, Massachusetts Institute of Technology, May 1984.
6. Bodewig, E., "Matrix Calculus", North-Holland Publishing Company, Amsterdam, pp. 125-142, 1956.

## Detection of Active Emitters using Triangulation and Trilateration Techniques: Theory and Practice

A.M.Dean  
GEC-Marconi Research Centre  
West Hanningfield Road  
Great Baddow  
Chelmsford  
Essex  
CM2 8HN  
ENGLAND

### 1 SUMMARY

Recent conflicts have highlighted the benefits of "soft-kill" electronic warfare (stand off, escort and self screening jamming), during intrusion into areas protected by Air Defence (AD) radar networks. These conflicts have highlighted the need to protect and supplement the Recognised Air Picture (RAP) with the ability to locate and track the intruding jammers. This paper describes a demonstrator Passive Jammer Location (P JL) system, and some of the theory behind it, currently under development at the Marconi Research Centre.

The two basic geometrical techniques for locating unknown emitters, usually termed *triangulation* and *trilateration*, are identified. The main problems associated with triangulation techniques, those of target ghosts and ghost resolution in denser scenarios, are discussed and trilateration processing using correlation offered as a solution.

The main feature of an operational P JL system is noted as being the need to positively resolve jammer positions, to sub-beam accuracy, in dense jamming scenarios. This includes the "pop-up" target appearing over the radar horizon and the agile sophisticated jammer. In addition a number of other features are identified which would be desirable in any future NATO P JL system.

Over a number of years the UK MoD and GEC-Marconi have undertaken a number of studies relating to P JL architectures and data processing techniques. In 1991 these studies led to the Air Defence Emitter Location Equipment (ADELE) Technology Demonstrator. The objectives of the ADELE programme are to demonstrate that the requirements of a P JL system can be met at a price affordable by AD system procurers. Additionally to confirm that the new P JL data processing techniques, developed during these previous studies, perform as predicted during live trials.

The paper discusses the main hardware and software modules making up the ADELE demonstrator including: the multibeam antenna, the resistive matrix beamformer, the P JL multi-channel signal sampling hardware, radar interfaces and synchronisation, signal and data processing, display and recording and simulation resources.

### 2 INTRODUCTION

We shall not attempt to discuss Passive Detection comprehensively in a single paper. Here we shall consider only the special case of Passive Jammer Location relating to airborne jammers in the Air Defence radar bands. Underlying this is the need to maintain a reliable and up-to-date Recognised Air Picture in order to commit defence resources efficiently. To allow sufficient response time, AD radar networks must have a medium/long range detection capability and must maintain surveillance over large volumes of airspace containing many aircraft. This type of capability is vulnerable to jamming, for reasons well known to students of the Radar Equation.

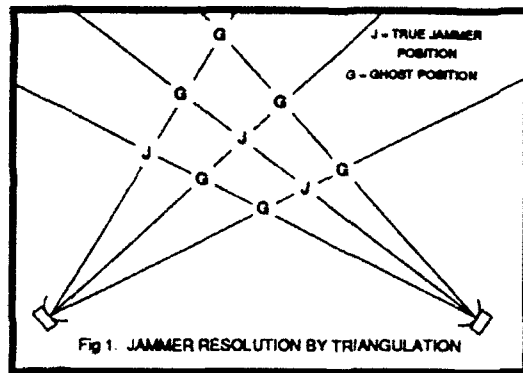
In contrast to predominantly "one on one" interactions between short range defence systems and their targets, jamming in AD radar networks typically involves many jammers. These will almost certainly include distant but very powerful Stand-Off Jammers (SOJs) attempting to dazzle the forward view of any AD radar within range. Groups of intruding hostile aircraft will carry sophisticated, agile jamming equipments which they will use when their on-board Radar Warning Receivers (RWRs) detect illumination by defence radars. To maintain the RAP the AD network must find, resolve and track such intrusions against the competing SOJ background. Also, if the network can report individual SOJ plots this opens the way to countering them by either physical or electronic means.

### 3 TRIANGULATION USING JAMMING STROBES

To be effective, jamming must be received at good strength by its victim radars. AD radars are normally fitted with means of determining the directions from which strong jamming signals are being received and these angular *strobe* reports, when combined from radars at different sites in the defence network, can be triangulated to find intersection points defining jammer locations within the surveyed airspace.

A widely-recognised problem, discussed in detail in the literature [1], is that strobe intersections are ambiguous. If each of two separated radars reports, say, 10 different strobe headings these could define

up to 100 different intersection points when triangulated. At best, not more than 10 of these can be genuine (but which?); at worst, in a low-elevation scenario typical of medium-range AD surveillance, none need be genuine since there could be 20 jammers, none simultaneously visible to both radars. This problem is shown pictorially in Figure 1 for 3 jammers. It can be seen that for  $n$  jammers, given ideal mutual visibility,  $n \times n$  intersections are produced of which  $n(n-1)$  are ghosts.



In practice, not all 100 intersections would be plausible and a number of logical and procedural tests can be applied to reduce the number of ghost locations. Jammer tracks can be tested for plausible speeds and manoeuvres and locations can be tested for consistent heights and existence in real space. These tests demand accurate strobe angle and angle rate reports both in azimuth and elevation. Unfortunately, in a typical low-level AD scenario all sources (both SOJs and intruders) will normally be seen close to the horizon where elevation is not a useful discriminant and where angular measurements tend to be corrupted by multipath.

Angular accuracy is easily corrupted in other ways too. Even the simplest type of angular interpolation, where the receiving beam scans steadily through an isolated source heading, relies on an assumption that the source strength remains steady. With agile, responsive jammers no such assumption can be made. In practice more advanced techniques, such as super-resolution [2,3] may be applied but these also depend on unsafe assumptions such as point sources, etc. These difficulties are magnified when groups of jammers are seen against an undifferentiated SOJ background: angular measurements are blurred and it becomes difficult to separate or even count strobes.

Simple strobe triangulation is thus of limited value against realistic jamming threats to AD networks. Its performance can be extended by applying advanced techniques (some of which are relatively costly) but, even so, it is likely to fail in the cases where it is needed most. Where RAP maintenance is vital, triangulation is inadequate and a more powerful PJL technique must be used.

#### 4 REQUIREMENTS FOR EFFECTIVE PJL

The ability to distinguish between jamming emissions from different sources underlies all aspects of PJL performance. Achieving this is difficult because jamming emissions are designed not to convey information but to deny it. Cooperating jammers may

radiate irregular overlapping emissions which cannot be resolved by tuning or by normal ESM techniques, in either the time or frequency domain.

Five features are essential for effective PJL in air defence radar networks:

- (i) Positive resolution of multiple jamming sources and strobes.
- (ii) Positive matching of jamming strobes received at different sites.
- (iii) Sub-beam plot accuracy for jamming sources.
- (iv) "Single-shot" detection and plotting of fleetingly-observed jammers.
- (v) A sensor architecture and management system able to "catch" brief, agile emissions "in-beam" at two sites.

Before presenting solutions to the problems we have posed, we must consider what other attributes are needed in a PJL system. While not all of these attributes are individually essential, taken together they exert a powerful influence in shaping PJL architecture and management. Without attempting to be comprehensive we can list:

- (vi) Automatic operation without skilled operators.
- (vii) Flexible compatibility with existing AD radars in the NATO inventory.
- (viii) No compromise to the operating autonomy of individual network radars.
- (ix) PJL reporting compatible with existing radar net reporting.
- (x) PJL mobility at least matching that of the radars being supported.
- (xi) Data links between sites to be simple, reliable, easily provided.
- (xii) Covert deployment and operation; PJL support not discernible by intercept.
- (xiii) Total PJL system costs only a small fraction of the radar network cost.

And possibly,

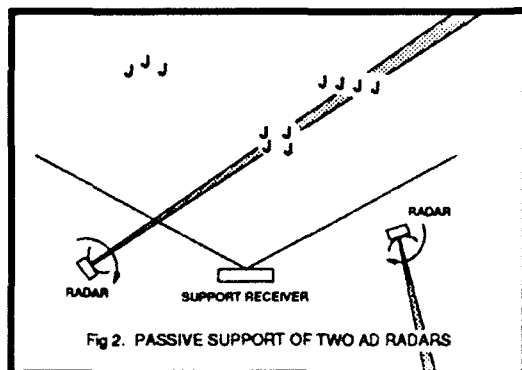
- (xiv) A simultaneous Active-Bistatic radar support mode.

None of these attributes will be contentious except perhaps the last, which therefore requires some comment. Active-bistatic radar, using a receiver distant from the transmitter, has long been known but has been little used in recent times because it is usually not a cost-effective alternative to a monostatic radar of comparable power and antenna size. However, bistatic reception by a distant receiver in addition to normal monostatic operation can provide important system performance enhancements [6] in stressful scenarios. The point of interest here is that PJL support, whether by triangulation or otherwise, exploits bistatic or multistatic dimensions within a radar network. A PJL system which can exploit these actively as well as passively has more to offer than one which cannot. Where this is done *simultaneously* using the *same* equipment, the additional cost can be very small.

## 5 SOLUTIONS FOR EFFECTIVE PJL

To satisfy the main requirements, noted above, special resolution techniques are required, such as cross-correlation [4,5] of signals received at different sites. Correlation can provide a high-resolution hyperbolic range coordinate for each jamming source, allowing sub-beam plot accuracy by trilateration. When this hyperbolic range is combined with the angular data provided by triangulation processing an accurate jammer position is identified. In practice correlation must be supported by the fullest use of conventional resources including narrow, high-quality sensor beams and good antenna sidelobe rejection [4]. Such requirements would make efficient PJL unaffordable were it not for the fact that most AD radars already have antennas with suitable narrow-beam characteristics. Affordable radar-band PJL *must* therefore exploit the antennas of the radars it supports.

Efficient PJL support within an AD radar network requires a *Passive Support Receiver* (PSR) in addition to the network radars themselves. A well designed PSR can support several radars simultaneously, provided of course that each radar has some useful field of view overlapping that of the PSR. A multi-beam PSR supporting two AD monostatic radars is shown in Figure 2. Within regions of overlap PJL support can meet all essential requirements and can have as many of the other attributes we have listed as the designer sees fit to provide.



To capture the matching jammer excerpts the PSR is managed so that it follows not the unknown jammers but its cooperating radar partners. The PSR turns its relatively broad sector of coverage in the general direction currently being illuminated by a radar requiring support and controls its tuning and capture timing to coincide with that of the radar. What the radar captures, perhaps unexpectedly, the PSR will also have captured automatically, so the two can be cross-matched. Provided only fairly short jamming samples need to be captured by each partner, the radar can maintain its autonomy in respect of scan patterns, operating modes and frequency schedules as long as it can provide information so that the PSR can follow it. For the same reason an agile PSR can support more than one radar simultaneously by interleaving capture timing.

A problem is immediately apparent, since a narrow PSR beam cannot cover a generous sector of view in the way we have assumed. Different solutions to this problem define alternative system architectures. Here we can only discuss two of these, both of considerable practical performance.

The first solution is to scan the PSR beam rapidly across a chosen sector of coverage during each dwell period of the cooperating radar. [4] This solution limits the maximum length of jamming samples captured from a particular target location, and requires a more dedicated degree of cooperative linkage than would ideally be desired. However, the scheme is entirely practical and can be engineered to give good results.

An even more powerful solution is to divide a generous PSR coverage sector into an appropriate number of *simultaneous* narrow sub-beams by using a modern beamforming technology of which we shall have more to say presently. When engineered to have appropriate bandwidth and frequency agility, this solution imposes little or no restriction on capture timing or on other aspects of cooperative linkage, which makes it easier for the PSR to support different types of AD radar together on the same network. Further, this type of PSR can more readily give simultaneous active-bistatic radar support since its synchronisation requirements for PJL and for active-bistatic support modes are essentially the same.

## 6 THE ADELE UK DEMONSTRATOR

Although the principles behind efficient PJL are well established, advances in technology allow these principles to be implemented in new ways, giving better performance at reduced cost. As the form of modern AD networks responds to changes in world security, PJL support remains as vital as ever but it must be more flexible. Effective support cannot be limited to the major quasi-static radars of national defence networks but must be free to follow current and future radars wherever they may be called to serve. Meanwhile, jammer technology continues to evolve; PJL systems must have performance margins and growth potential to counter threat developments which cannot yet be foreseen.

With considerations such as these in mind, the Defence Research Agency, in cooperation with GEC-Marconi and in consultation with NATO through SHAPE Technical Centre, have initiated the ADELE PJL Technology Demonstrator Programme.

The aim of the ADELE programme is to demonstrate:

- ø That all features and requirements relevant to the passive support of AD radar networks can be combined into a compact PSR unit which can be deployed as required.
- ø That new performance goals, based on advanced types of processing, are achieved in practice under realistic trials conditions.
- ø That effective passive support can be provided at affordable cost.

The ADELE Demonstrator is based on the BEARS multibeam antenna developed cooperatively by GEC-Marconi and UK MoD(PE). The BEARS antenna, currently sited at the GEC-Marconi Research Centre near Chelmsford, UK, has already been demonstrated giving active-bistatic support to the MoD BYSON experimental AD radar sited at the Defence Research Agency, Great Malvern, some 200 km distant. This work has been reported elsewhere [6,7]. The ADELE programme complements this by demonstrating the same antenna performing a PJL support role.

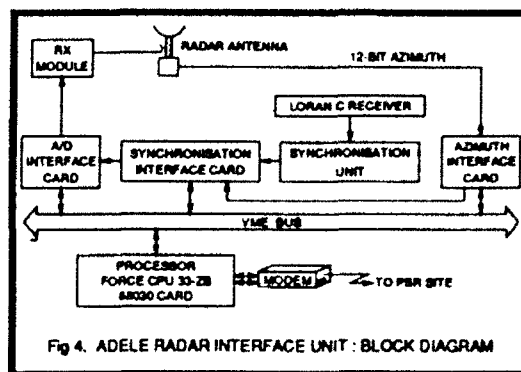
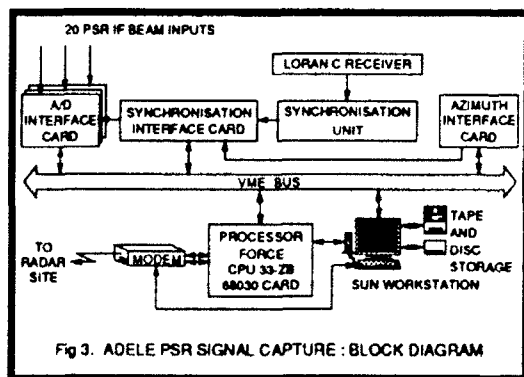
Initial local PJL trials using this antenna were successfully completed in the late Summer of 1991 and the equipment is currently being prepared for larger scale trials later this year in which the antenna will be linked to a radar sited about 10 km away. Substantive UK trials over long baselines are scheduled to take place during 1993/94 and a 250 km International Trials linkage spanning the North Sea is provisionally planned for 1994.

## 7 THE KEY TECHNOLOGIES

In the remainder of the paper we will look briefly at some of the key modern technologies lying at the heart of the ADELE PJL demonstrator. Leaving aside one or two aspects which we cannot discuss, these are:

- ø The BEARS antenna architecture.
- ø The simultaneous multiple beamformer.
- ø Multi-channel signal sample recording.
- ø Interfaces with cooperating radars.
- ø Synchronisation between sites.
- ø Signal and data processing.
- ø Display and recording.
- ø Simulation resources.

Figures 3 and 4 show block diagrams of the PSR Signal Capture and ADELE demonstrator Radar Interface Units respectively.



### 7.1 Antenna Architecture

The antenna must form narrow radar-quality beams but must be compact and covert. There is no transmitter so an efficient receive-only antenna architecture can be used. The BEARS antenna is a flat phased array composed of thin vertical "planks" populated with Woloszczuk slot dipoles [8]. Each plank carries its own triplate vertical distribution and its own hybrid receiver front end module. The very modest front-to-back depth makes it easy to mount the planar assembly conformally in the side of a small motor van or transport container.

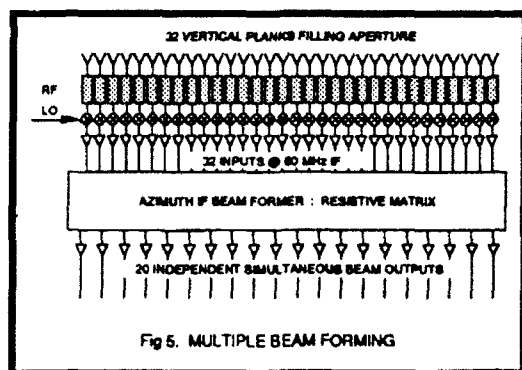
The vertical distribution defines the elevation beam pattern. For use in medium to long range AD networks the elevation pattern peaks just above the horizon and tapers off at higher angles. The azimuth pattern is determined by the number of planks used, and by the azimuth beamformer which is quite separate. The BEARS antenna normally uses 32 of the 36 planks installed. This gives about 2.5 sqm of aperture, which is sufficient for reasonable active-bistatic radar sensitivity, and allows 4-degree azimuth beams to be formed in the 3 GHz radar band. The number of planks installed and used could be varied to suit space and performance requirements. An operational version of the BEARS antenna might well have more than 36 planks installed.

### 7.2 Multiple Beamforming

Since the PSR has no transmitter, and since the receiver noise figure is determined by the front-end modules in each plank, there is no penalty for implementing beamforming at intermediate frequency. Alternative analogue or digital beamforming technologies can be used, simultaneously if necessary. For example, PJL and active-bistatic reception can be performed simultaneously using different beam patterns. Here we shall describe only the multibeam arrangements normally used for PJL operation, although other options, including an adaptive beamformer, are in fact installed.

The BEARS PJL IF beamformer, shown in block diagram form in Figure 5, operates at 60 MHz centre frequency and consists of a film-technology resistive matrix having 32 plank inputs and 20 simultaneous beam outputs. The fan of 20 azimuth beams is designed to cover a 90-degree sector at mid-band, but any other number of beams, or pattern, could be generated by adapting the resistive matrix appropriately.





The beamformer supports an instantaneous signal bandwidth in excess of 20 MHz, which is more than sufficient to match the bandwidth of captured jamming samples relayed from cooperating radars. Adjustment of the receiver local oscillator (LO) frequency, which is common to all the receiver front ends, tunes reception to any desired frequency within the design range of the aperture, without affecting operation of the beamformer. The system is designed so that agile control of the LO will allow the PSR to follow frequency-agile radars.

### 7.3 Multi-Channel Signal Capture

Jamming samples received in the 20 BEARS beams must be captured together if all are to overlap in time with a sample relayed from a cooperating radar. Capture must, of course, be implemented in real time, whereas temporary storage of signal samples and cross-processing are most economically implemented by a digital system able to deal with beams serially rather than all together. In the Demonstrator the real time/machine time gap has been bridged by providing a special analogue/digital interface card for each beam channel in which signal samples are to be captured.

The inputs to these cards connect directly to the beamformer outputs of the BEARS antenna, at 60 MHz IF, and each card carries out all the signal conditioning required to prepare digitised baseband signals for capture. Operation is continuous but when gated by a synchronised signal capture command pulse, each card "freezes" a suitable length of signal excerpt in its digital memory. These excerpts are then downloaded serially into the digital processing system.

### 7.4 Radar Interfaces

Signal excerpt samples, beam heading, timing information, and other data must be extracted from cooperating radars and relayed to the PSR. For experimental ADELE trials it is also necessary to relay certain control and diagnostic data not needed in an operational system. A special interface unit is provided in order to equip existing AD radars to provide samples and data in the form required. This is coupled to a cooperating radar and to an intersite link. For the Demonstrator trials this link will normally be national or international telephone lines.

Capture of jamming signal samples from a cooperating radar can be implemented at RF, IF or baseband video according to the beamforming and receiving channel architecture used by the radar. For the Demonstrator this is done at RF and IF, by providing a duplicate BEARS-type receiver front end module coupled to an appropriate radar beam port, and by providing a duplicate IF signal capture A/D interface card within the interface unit. The unit also receives 12-bit azimuth data from the radar which it relays to the PSR, together with timing information.

### 7.5 Synchronisation

Provided jamming signal excerpts captured at different sites overlap in time sufficiently to allow them to be cross-processed, exactly simultaneous capture is not essential. However, very accurate knowledge of the relative timing is necessary in order to extract accurate range information for any targets which may be found. This applies both to PJL and to active-bistatic radar detections.

The requirement is met by having accurately synchronised clocks at each site. Recorded jamming excerpts are time-tagged at each site so that when subsequently cross-processed their relative timing is known to the accuracy required for plot reports. The compact atomic clocks need to be resynchronised only at infrequent intervals. The ADELE Demonstrator uses Loran C equipment for this purpose.

### 7.6 Signal and Data Processing

A multibeam PSR having good instantaneous bandwidth can capture a great deal of signal data which must be sifted and processed or discarded while still fresh. The only economic approach is to use industry-standard digital processing hardware for all signal and data handling subsequent to the special A/D channel interface cards we have already mentioned. The exclusive use of standard digital hardware distinguishes the ADELE PJL system from its predecessors and is one of the features which the Demonstrator programme seeks to assess.

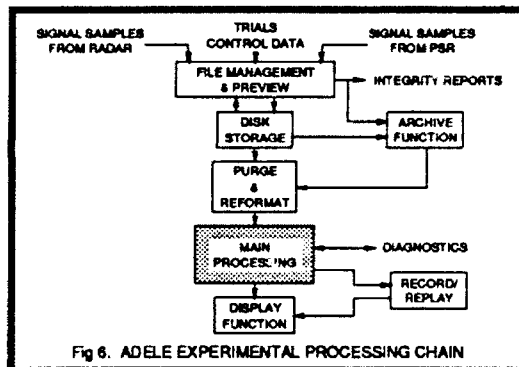
An advantage of non-dedicated hardware is that it can support alternative and adaptive forms of processing. For the Demonstrator it means that alternative processing strategies and algorithms can be tested and the results critically compared. This requirement is reflected in the processing resources provided for the ADELE Demonstrator.

The Demonstrator processing architecture is designed so that all captured signal samples, in all beams, can be recorded and stored for repeated replay through alternative processing. Although this would not be a feature of an operational PJL system, for an experimental Demonstrator it has several advantages. Firstly, it enables the maximum possible value to be obtained from signal samples captured during multiple-jammer trials and Air Defence exercises which, for obvious reasons, cannot be repeated at will to suit processing experiments. Secondly, because it is not essential to complete the processing sequence in real time, it allows existing shared processing resources to be exploited. Finally, it permits diagnostic intrusions into the processing sequence in order to analyse the efficiency of algorithms in as much detail as may prove necessary.

The interface unit fitted to the cooperating radar uses a Force CPU 33-ZB single-board computer card coupled to a Motorola Codex 3385 modem. Associated with the 68030/25MHz CPU are 4MB of RAM, two serial I/O channels, timers, and a real-time clock. The modem communicates at up to 19.2 kbaud, which is sufficient to relay captured jamming strobe samples but not for the real-time relay of signals received at all beam headings when recording for subsequent replay. Two operating modes are therefore provided. In "recording" mode all captured signals are stored in full for a limited observation period, and then relayed. In "real-time" mode only "important" captured signal samples are relayed, as in real life, but continuous operation can be maintained for extended periods.

A similar single-board computer and modem are used at the BEARS multibeam site but these are supplemented by a SUN IPX workstation interfacing with disk storage for captured signal data awaiting processing, and tape storage for longer-term retention of trials "runs". The interface unit at the cooperating radar site can be initialised and controlled remotely from the SUN workstation via the modems and link. Software for these systems is written in "C" language.

The main software routines and data flows used during the trials phase are indicated in Figure 6 and described below.



During trials "runs" in which signal data is being recorded at high rates for later processing and analysis it is desirable to have confirmation that the recorded data is not corrupted in any way, that its area of coverage is adequate, and that it does include samples from all important jamming sources. This information is required very promptly so that, if necessary, adjustments can be made while airborne test targets remain in view (there are no "targets of opportunity" for PJL in peacetime). The ADELE Demonstrator therefore includes a special processing function which previews captured signals as they are stored and reports their integrity and content.

Extended trials will generate large numbers of stored data files. These are coordinated by a file management program to ensure that stored data is identified unambiguously, remains linked to associated data such as the trials log, and can be recovered at will for replay and processing. Where full signal samples have been recorded in all beams the processing load can be reduced by rejecting samples of no interest. Many of the processing algorithms, along with extensive simulation resources, run in FORTRAN 77 which is the preferred medium for non-real-time processing. The

interface between storage and processing therefore purges and reformats the stored data. These facilities relate to the experimental status of the Demonstrator and will be absent in an operational version which will process or discard all signals within a few seconds of capture.

The suite of PJL detection processing and plot extraction algorithms is designed to run in any type of general-purpose processing hardware. During the development phase, detection processing is performed on a general-purpose multiple-user network of SUN IPX workstations installed at the GEC-Marconi Research Laboratories.

In its present form the PJL detection processing suite is generously equipped with subroutines which provide diagnostic data outputs for analysing the performance of alternative algorithms and processes. For trials use these will be bypassed in order to improve the overall processing response time. Some latency is inevitable in any multistatic system which adaptively processes selected jamming samples; normally this latency would not exceed the antenna scan periods of the network radars being supported.

### 7.7 Display and Recording

Operational PJL systems must report jammer plots externally using appropriate formats but the ADELE Demonstrator will have a local PPI display using an industry-standard colour monitor. There will also be relay facilities for remote display, for example at the site of a cooperating radar during trials. At the BEARS site the display shares facilities already installed for active-bistatic trials but with the supporting software revised to provide additional features. These features include:

- Hyperbolic trilateration curves relating to location by cross-correlation.
- Symbolic target representation indicating auxiliary detection parameters.
- Operator interaction to display additional data for selected targets.
- Zoom expansion for target groups.
- Artificial persistence to simulate raw target tracks.

The display facilities are equipped for recording the data inputs in full during trials runs. This enables runs to be replayed at increased or reduced frame rates, and allows the operator to "freeze" individual frames and to call up and examine the full detection parameters for targets of interest.

### 7.8 PJL Simulation

A realistic PJL test scenario must include a substantial number of different airborne jammers. It is practical to provide this only rarely, during major air defence exercises. Occasional large-scale PJL trials are essential but they cannot adequately test the full range of latent performance, so "real" trials must be supplemented by simulations. Indeed, simulation is so important in PJL work that we might claim that it is the "real" trials that are supplementary.

It is obvious that comprehensive testing by simulation must embrace multi-source jamming strategies spanning realistic AD scenarios out to long ranges. What is less obvious is the amount of fine detail necessary to represent individual jammers and their emissions, because real world jammers do not emit neat mathematical waveforms. It is necessary to model interactions between multiple emissions and receiving channel characteristics, including phase-sensitive phenomena, in order to model detection and resolution performance with good quantitative accuracy.

Simulation resources having these and other capabilities have been built up at the GEC-Marconi Research Centre. The simulations run in FORTRAN 77 on SUN IPX and VAX networks. The simulations generate captured jamming signal samples in formats identical to those used by the real PSR and associated cooperating radars and are designed for direct connection to the processing resources used for detection processing during actual trials.

Comprehensive simulation resources of this class constitute an essential working tool without which today's advanced adaptive PJL processing algorithms could not have been developed and proved. It is clearly essential that the simulations validly represent the characteristics of real-world emissions and receivers so, at each successive stage of the ADELE trials programme, even greater importance is placed on demonstrating the accuracy with which the simulations model the actual trials results than on the excellence of the results themselves.

## 8 CONCLUSION

The importance of radar surveillance to air defence strategy encourages an aggressor to suppress defence radar performance wherever possible. Jamming is a cheap but powerful tool for this purpose. The ability to track airborne jammers accurately is of immediate value to the defence and a disincentive to the aggressor.

Where AD radars are networked, isolated jammers can be located by strobe triangulation when conditions are favourable but this method has fatal weaknesses in denser scenarios. Under these denser conditions very much more powerful processing methods based on correlation are considered essential. These methods exploit jammer reception by the radars themselves but require an additional Passive Support Receiver.

Until comparatively recently, really effective PJL support has been possible only in major networks. This has been revolutionised by new technologies which make efficient passive support an affordable "Add on" option for AD networks everywhere.

The key technologies include compact frequency-agile antenna modules which permit large arrays of simultaneous staring beams to be formed cheaply at intermediate frequency. A card has been developed which extracts jamming signal samples from radar and PSR beams and presents the samples in a form acceptable to general-purpose industry-standard digital processors. An add-on interface unit has been developed to convert existing radars for PJL support.

Digital memory can hold large arrays of captured signal samples for long enough to identify those of real importance and to relay them via low-bandwidth links to distant sites. The very high processing capacity now available in compact low-cost machines supports advanced adaptive processing which can respond to future developments in the jammer threat. PJL processing can share machine capacity with other system functions according to demand and priority.

These technologies enable an architecture in which one passive receiver can support several netted radars and which is easy to manage and synchronise. The multibeam unit can be compact, covert and mobile and can provide active-bistatic radar plots as well as jammer location when linked to network radars by telephone-quality links.

The need for effective passive support in the AD radar bands and the principles by which it is achieved have been proved over 30 years; these are not matters of doubt. That this support can now be given throughout NATO at low cost and at a new level of efficiency will be shown in forthcoming trials of the ADELE Demonstrator.

## 9 ACKNOWLEDGEMENTS

This work has been carried out with the support of the Defence Research Agency.

## 10 REFERENCES

1. "Passive Jammer Location by Trilateration" : G.N.Taylor IEE Colloquium : Signal Processing for Electronic Warfare : January 1992
2. "Superresolution using an Active Antenna Array" : U.Nickel Proc. IEE : "RADAR 82" : pp. 89-91 : October 1982
3. "Direction Finding with Linear Phased Arrays using Eigen-Analysis Techniques" : P.G.Starkey GEC Journal of Research, Vol. 5, No. 4, pp. 193-207 : 1987
4. "Mixed Triangulation/Trilateration Techniques for Emitter Location" : F.J.Berle Proc. IEE : Pt.F, Vol. 133, No. 7, pp. 638-641 : December 1986
5. "The ECCM Potential of a Passive Bistatic Complement to Radar Detection of Aircraft Targets" : D.Wooler & A.J.Poelman AGARD Symposium on Avionic Countermeasures : October 1990
6. "Bistatic Radars for Air Defence" : M.R.B.Dunsmore Proc. IEE : "RADAR 87" : pp. 7-11 : 1987
7. "Description of an Experimental Bistatic Radar System" : T.A.Soame & D.M.Gould Proc. IEE : "RADAR 87" : pp. 12-16 : 1987
8. UK Patent No. 8523076/18:9:85

# RADIOLOCATION OF A SATELLITE-BORNE LOVHF BEACON

Robert B. Rose  
 Ionospheric Branch, Code 542  
 Ocean and Atmospheric Sciences Division  
 Naval Command, Control Ocean Surveillance Center, RDT&E Division  
 San Diego, CA, 92152-5000, USA

## ABSTRACT

A year long experimental program was conducted to measure refractive bending, or how much the signal deviates from true line of sight, at low VHF frequencies (29.5 MHz), and to determine whether this deviation or error could be predicted using large scale ionospheric models such as the Ionospheric Conductivity and Electron Density (ICED) program. An experiment to directly measure the angle of arrival of a 29.5 MHz signal from an orbiting satellite was successfully completed. The satellite was in a circular orbit at an altitude of 1000 km. It was shown that refractive errors can be directly related to the electron density along the measurement slant range. Ionospheric disturbances such as sporadic E and ionospheric storms produce large, short term errors that can approach 10 degrees. In addition to day/night variations, seasonal and solar cycle sensitivities were found. The refractive error varied so rapidly with respect to time and space that its prediction with a median value ionospheric code is almost impossible.

## INTRODUCTION

Recent advances in the science of ionospheric measurement provide a detailed picture of how the ionosphere is structured with respect to time and space. What is seen is a much more turbulent plasma than previously recognized, one that is very susceptible to changes in solar radiation patterns and the earth's magnetic field. This new knowledge raises questions about the predictability of ionospheric structure. The bulk of the models used to predict electron density between 50 km and 1000 km are based on empirical measurements. Their outputs are median values. The key concern is how well they typify the real-time ionosphere as a function of time of day, season and geographical location.

Programs at this facility use a variety of ionospheric models. Usually, the first question asked is how well they can replicate the dynamics of the real world ionosphere. Recently this question was posed about the Ionospheric Conductivity and Electron Density (ICED) model (Daniell, et al., 1986). The application of interest required an accurate prediction of the electron density between the ground and outer space. One output from ICED is a true height electron density profile between 50 and 1000 km. A predicted total electron content (TEC) was derived from this. The

challenge was to design an experiment that would measure TEC, or something equivalent, to test ICED model profiles.

The discovery of an orbiting satellite system that carried a low VHF (LoVHF) beacon operating at 29.5 MHz suggested a novel approach to this problem. The remainder of this paper describes an experimental program in which a ground based HF interferometer was used to locate and track the LoVHF beacon as it passed overhead. The objective of this experiment was to (1) directly measure the refractive error, or how much the signal deviates from true line of sight, on the transionospheric signal; (2) determine the magnitude and characteristics of this uncorrected error; and (3) develop an electron density data base against which the ICED model could be compared.

## EXPERIMENTAL METHOD

The signal source for this experiment was a low orbiting satellite. It contained several HF/LoVHF beacon and transponders which produced morse code signals at 29.5 MHz. For this project, the morse beacons were used as point signal sources. Originally planned as two individually orbiting platforms, the satellites were launched together on a navigational satellite (NAVSAT). This was fortunate as the NAVSAT was maintained in an accurate circular orbit at 1000 kilometer altitude. This encompasses 98 percent of the ionosphere. The satellite's orbital inclination of 83 degrees provided two or three useful passes a day. Data were collected every 15 seconds as long as the 29.5 MHz beacon signal strength was sufficient for gathering data. A normal pass would last 12-13 minutes from horizon to horizon.

Conceptually, the experiment was simple. The "apparent" position of the satellite was measured and compared to the "real" position derived from predicted ephemeris data specifying the orbital location of the satellite in space and time. The difference, expressed in angular error in degrees, is the refractive bending error caused by the ionosphere between the satellite and the ground station. The electron density was then derived using coefficients of refraction at the beacon frequency.

To measure the satellite-borne beacons, a direction finder with fast processing time was required. It had to make high time-resolution measurements of angle of arrival (AOA) in azimuth and elevation. At that time, the only system capable of accomplishing this was a Single Site Locating (SSL) Testbed at Southwest

Research Institute (SWRI) in San Antonio, Texas. The system, described earlier (Rose, 1992), is a 7 element interferometer which uses an "L" shaped array. It can perform successive angle of arrival measurements in 3.5 millisecond intervals. At this sampling rate errors due to satellite motion are effectively minimized.

In order to perform the desired comparisons with ionospheric models, the refractive bending data had to be expressed in terms of electron density. Before and after each pass, the peak electron density was measured with a vertical incidence sounder. The measured F-region critical frequency, the foF2, is directly related to peak electron density (Ne) by:

$$(1) \text{ Ne(el/cm}^3\text{)} = 1.24 \times 10^4 \times \text{foF2(MHz)}$$

Because the interferometer produced large amounts of data very quickly, data collection periods were limited to windows that were four seconds in length. In this time 1152 frames of data were collected. A frame is a 3.5 millisecond AOA measurement of the phase data from each of the seven interferometer elements. The satellite moved approximately 30 km during the 4 second window or about 26 meters between frames. Each frame was time tagged and stored on magnetic media. In addition, a vertical incidence sounder ionogram was made prior to and after each pass to determine ionospheric conditions. Next, each frame underwent phase linearity testing. This test determines whether the array observed a plane-wave and made an AOA observation based on one signal. The frames that passed this test, typically 100-500 of the 1152 collected in each window, were put into an analysis file for further processing. Figure 1 shows an analysis file segment from one four second window. In a final step, the data are reduced into spreadsheet format for analysis.

#### DISCUSSION OF THE DATA

The experimental program was conducted between October 1988 and December 1989. Over this period, 217 passes provided information that were suitable for analysis. These included:

- 28 Cases - Setup/Calibration - (October - December 1988)
- 40 Cases - Winter- (January - March 1989)
- 40 Cases - Spring (April - June 1989)
- 50 Cases - Summer - (July - September 1989)
- 52 Cases - Fall (October - December 1989)

This experiment performed above expectations throughout the year. There was no period in which data could not be acquired.

A unique aspect of this experiment was the amount of the ionosphere that was measured on each pass. Depending on how

TIME HH:MM:SS.SS	TRUE AZ EL	SLANT RANGE	OBSERVED AZ EL	DELTA AZ EL
03:55:46.00	47.2 42.4	1365	47.5 42.9	-0.3 -0.5
03:55:46.01	47.2 42.4	1365	47.7 42.6	-0.5 -0.2
03:55:46.02	47.2 42.4	1365	47.2 42.9	0.0 -0.5
03:55:46.03	47.2 42.4	1365	48.1 42.7	-0.9 -0.3
03:55:46.04	47.2 42.4	1365	47.9 42.5	-0.7 -0.1
03:55:46.05	47.2 42.4	1365	47.3 42.7	-0.1 -0.3
03:55:46.06	47.2 42.4	1365	47.5 42.4	-0.3 0.0
03:55:46.08	47.2 42.4	1365	48.0 42.7	-0.8 -0.3
03:55:46.09	47.2 42.4	1365	47.6 42.8	-0.4 -0.4
03:55:46.10	47.2 42.4	1365	47.4 43.2	-0.2 -0.8
03:55:46.12	47.2 42.4	1365	46.9 42.6	0.3 -0.2
03:55:46.30	47.1 42.3	1366	47.0 42.9	0.1 -0.6
03:55:46.31	47.1 42.3	1366	47.4 43.0	-0.3 -0.7
03:55:46.33	47.1 42.3	1366	47.3 42.7	-0.2 -0.4
03:55:46.34	47.1 42.3	1366	47.4 42.6	-0.3 -0.3
03:55:46.35	47.1 42.3	1366	47.4 42.6	-0.3 -0.3
03:55:46.36	47.1 42.3	1366	47.6 42.8	-0.5 -0.5
03:55:46.37	47.1 42.3	1366	47.6 42.7	-0.5 -0.4
03:55:46.38	47.1 42.3	1366	47.5 42.4	-0.4 -0.1
03:55:46.39	47.1 42.3	1366	47.2 42.8	-0.1 -0.5
03:55:46.42	47.1 42.3	1366	47.0 42.6	0.1 -0.3
03:55:46.43	47.1 42.3	1366	47.6 42.8	-0.5 -0.5
03:55:46.44	47.1 42.3	1366	47.2 42.5	-0.1 -0.2
03:55:46.53	47.1 42.3	1367	47.0 42.8	0.1 -0.5
03:55:46.54	47.1 42.3	1367	47.0 42.8	0.1 -0.5
03:55:46.55	47.1 42.3	1367	47.7 42.7	-0.6 -0.4
03:55:46.56	47.1 42.3	1367	47.6 42.9	-0.5 -0.6
03:55:46.57	47.0 42.3	1367	47.7 42.5	-0.7 -0.2
03:55:46.58	47.0 42.3	1367	47.3 42.8	-0.3 -0.5
03:55:46.59	47.0 42.3	1367	47.2 42.8	-0.2 -0.5
03:55:46.60	47.0 42.3	1367	47.1 42.9	-0.1 0.3
03:55:46.61	47.0 42.3	1367	47.2 42.9	-0.2 -0.6
03:55:46.63	47.0 42.3	1367	47.2 42.9	-0.2 -0.6
03:55:46.64	47.0 42.3	1367	47.7 43.0	-0.7 -0.7
03:55:46.65	47.0 42.3	1367	46.9 42.6	0.1 -0.3
03:55:46.66	47.0 42.3	1367	47.0 42.8	0.0 -0.5
03:55:46.67	47.0 42.3	1367	47.4 42.8	-0.4 -0.5
03:55:46.73	47.0 42.3	1367	47.3 42.4	-0.3 -0.1
03:55:46.75	47.0 42.3	1367	47.9 42.6	-0.9 -0.3
03:55:46.77	47.0 42.3	1367	47.0 42.7	0.0 -0.4
03:55:46.78	47.0 42.3	1367	47.2 43.0	-0.2 -0.7
03:55:46.80	47.0 42.3	1368	47.1 42.7	-0.1 -0.4

Figure 1 - Segment of phase linear data from a four second frame

close the satellite passed overhead to the receiver, the data mapped represented an area about 600 km wide and several thousand kilometers long. The predicted ephemeris data were checked on each pass by carefully observing the times of acquisition of signal (AOS) and loss of signal (LOS). This normally occurs as the satellite comes over the horizon and is in line of sight of the receiving station. Any error in the predicted ephemeris data showed up as a bias in the output data and was easily spotted. AOS on a north to south pass was the most reliable test with the difference between predicted and observed AOS times never exceeding several seconds. Prediction of LOS to the south was completely unreliable, with the differences between predicted and observed being tens of seconds to minutes. This was because the 29.5 MHz signal would become a skywave signal at low elevation angles and could be heard half way around the world.

Most of the data were collected between elevation angles of 20 and 82 degrees. Below 20 degrees, the slant ranges exceed 2000 km, the signal to noise is very poor and the likelihood of an undistorted plane wave signal was very low. Directly overhead above 82 degrees, the arctangent calculation in the azimuth calculation breaks down and answers are unreliable. Once these boundaries were established, data collection became straight forward.

Several trends appeared early in the tests and it soon became easy for the

analyst to discern whether the pass was north to south or vice versa. The large patches of irregularities that characterize latitudes below 30 degrees north were evident. The ionosphere to the north of the receiver site was less dense than to the south. Changes in signals from the north were orderly and at expected levels. Signals from the south were always variable and unpredictable with greater errors at long slant ranges.

Because of the satellite's orbit, the times signals could be received changed daily. This allowed a measurement schedule that could probe the ionosphere at different times of day and night. The periods at pre-sunrise (electron density minimum) and pre-sunset (electron density maximum) were always the most interesting. At pre-sunrise in the winter, the electron density almost disappears allowing the scientist to determine the baseline accuracy of the interferometer without interference from the ionosphere.

Figure 2(a) shows a pass that occurred during the winter at pre-sunrise (absolute minimum electron density). To the north, there is little error in either azimuth or elevation angles that are measured. For this pass, the overall miss-distance error between predicted and observed AOA is 1.5%. Numerous other passes that occurred during the pre-sunrise period show about the same accuracy. Therefore, this is close to the minimum error of the SSL.

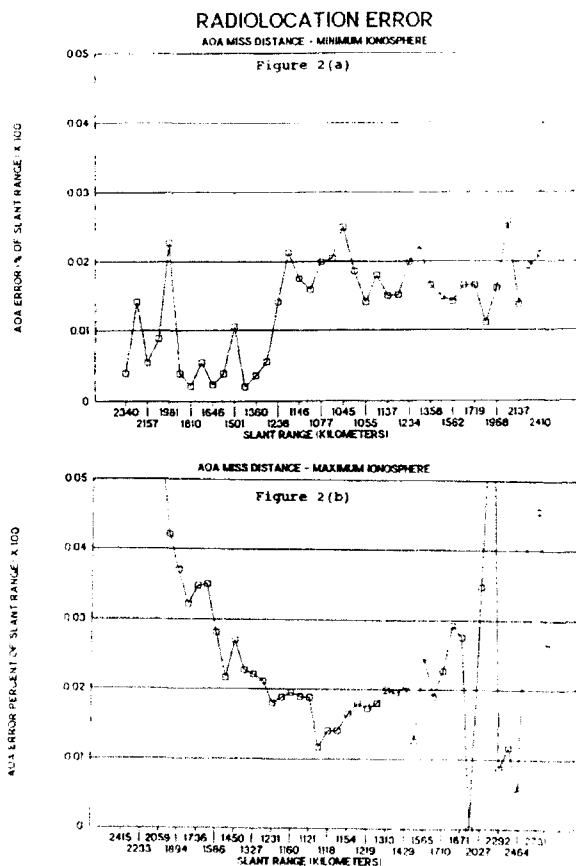


Figure 2(b) shows a summer pre-sunset case. The refractive error is largely effected by slant range, increasing as ranges increase. The average overall AOA error is 3.5%. However, as range increases, the error exceeds 5%, similar to the accuracies of the SSL in locating terrestrial targets.

Figure 3 shows the effects of seasonal changes and how refractive error changes with season. Figure 3(a) depicts winter night where error is minimal due to lower electron density. Even at night there is some evidence of the differences in the ionosphere to the north and the south. Figure 3(b) shows what happens during the night in the summer when electron density is greater. In both cases the data were collected during pre-sunrise when electron density is lowest.

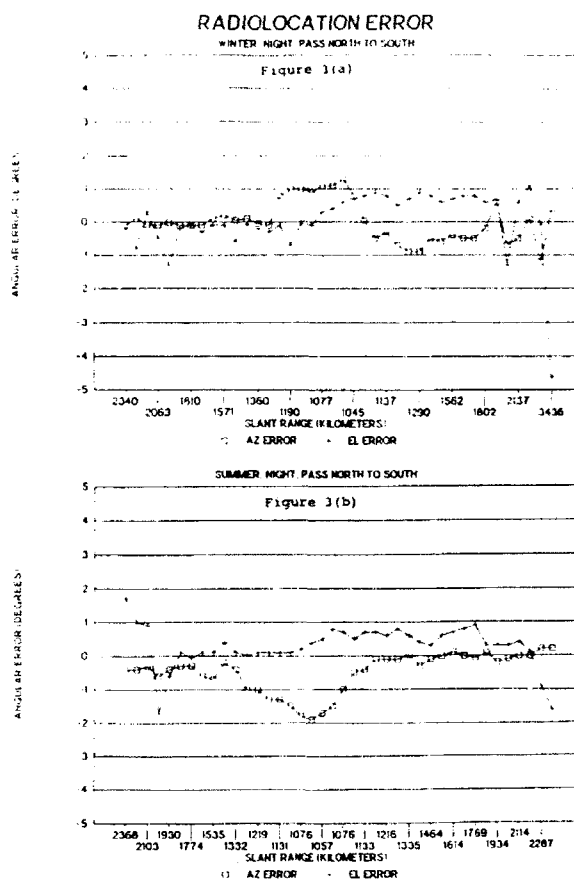
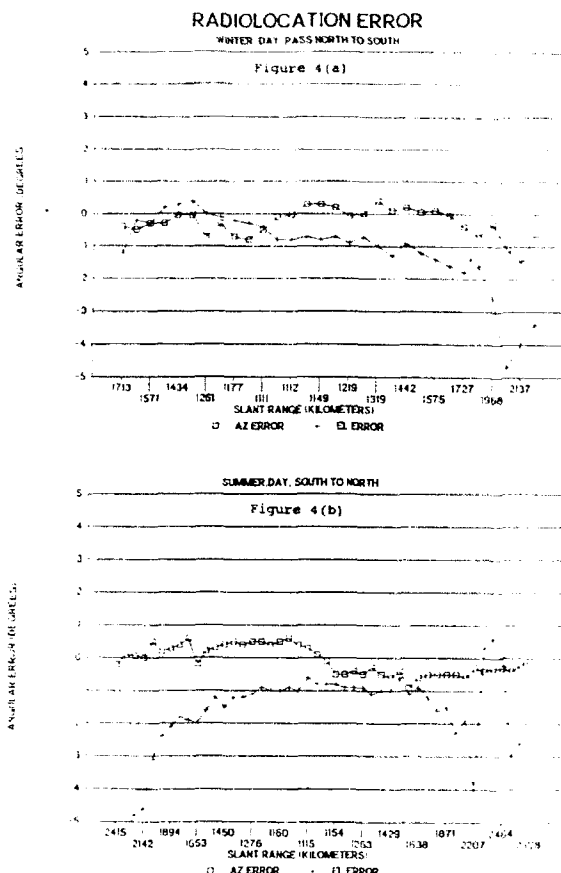


Figure 4 shows the effect of season on refractive error during daylight when electron density is at maximum values. The largest difference is in elevation angle error. In figure 4(a), the winter day, there is little refractive bending until the signal source is well south of the receiver site, and the errors are less than one degree. In figure 4(b), the summer day, there is a more pronounced signal bending over the longer slant ranges. For both examples the point of closest approach is nearly the same and at that point the errors are nearly the same.



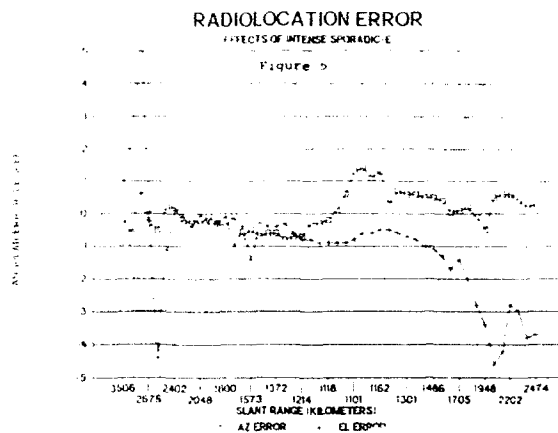
Analysis of the individual spreadsheet data for each pass showed that the errors in the elevation angle were the primary contributor to the overall error. Elevation errors are caused by horizontal electron density gradients which, in turn, are caused by cyclic changes in time of day, season, and solar cycle. These changes are smoother and less radical than those caused by vertical gradients.

Vertical gradients cause errors in the azimuth angle of arrival. These tend to be caused by short term phenomena such as sporadic-E, traveling ionospheric disturbances and ionospheric storm effects. Figure 5 shows the effect of sporadic E on a summer morning pass. The positive shift in azimuth error when the satellite is nearest the receiver (1101 km) is due to the very intense overhead E region. As the signal source moves away, the effects diminish rapidly because the layer is relatively thin.

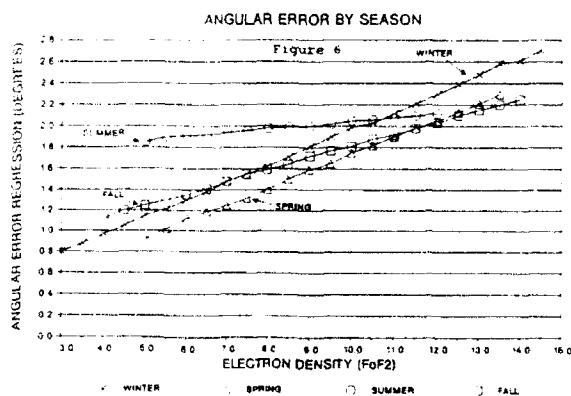
After a year of making measurements, there was sufficient data to observe the diurnal and seasonal ionospheric effects on AOA measurements. Each pass was reduced to a composite vector sum of the elevation and azimuth errors as a function of foF2. This was then plotted by season (Figure 6). While the results shown in figure 6 are not surprising, it does provide a quantitative relationship that can be used to test computer simulations.

The deviation of the summer curve from the other three seasons is referred to as the "seasonal anomaly."

The curve in figure 6 indicates that the relationship between transionospheric angular error and electron density is nearly linear. They can be expressed with two terms, one for winter, spring and fall, and one for summer. The measurements were made at solar cycle maximum when electron densities are the highest. Therefore, the relationships for solar minimum are already accounted for because it is already known that one of the effects of solar decline is that foF2 values are less as depicted in the lower half of figure 6. The data in figure 6 represent the "slowly varying" component for modeling transionospheric electron density and the resultant refractive bending.



#### TRANSIONOSPHERIC EXPERIMENT 29.5 (MHZ)



With data acquisition complete, the next objective was to compare observed data with predicted data from ICED. For the exact times measurements were made, model ionospheres were derived from ICED. The homing mode of a raytrace program was used to trace the signal ray path from the satellite to some ground point. Ephemeris data on the satellite's location were used as a starting point. The output was the latitude and longitude of the ground intercept point. This was converted into azimuth and elevation error and then

compared to the measured data. One method of displaying the comparisons is shown in figure 7. The differences between observed and predicted azimuth and elevation errors are used to generate a differential azimuth/elevation scatter plot. If the model predictions are close to the actual measured data, then the scatter plot provides a tight cluster about zero. The more the prediction deviates from the observed data, the greater the spreading of the data. Figure 7 shows data comparisons for both day and night situations between elevation angles of 30 and 60 degrees, the optimum angles for the interferometer array.

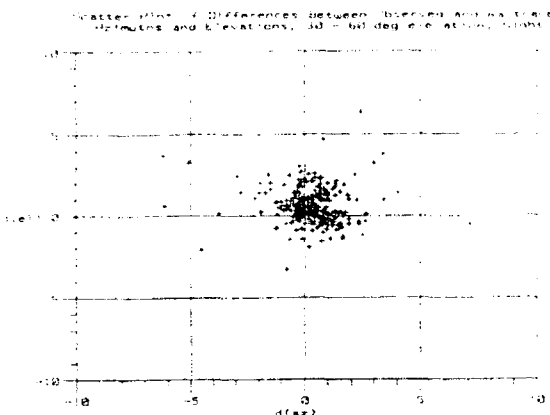


Figure 7(a)

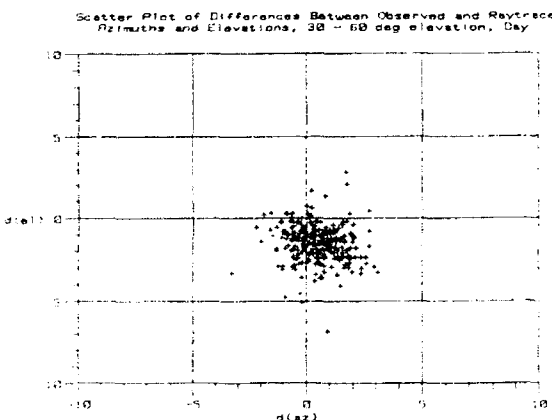


Figure 7(b)

Figure 7(a) represents 246 nighttime data points for true elevation between 30 and 60 degrees. The sample mean elevation difference value is 0.6 degrees with a standard deviation of 1.15 degrees. The mean azimuth difference was 0.25 degrees with a standard deviation of 1.36 degrees. With the minimal night ionosphere the predicted data set does track the observed data set, although with a tendency to under-predict elevation.

In figure 7(b), 269 daytime data points are shown. A fairly tight distribution is seen but the center of the elevation distribution has become negative, indicative of over-prediction. The mean elevation difference is -1.54 degrees with a standard deviation of 1.12

degrees. The mean azimuth error is 0.45 degrees with a 1.0 degree standard deviation. In combination with the results for the nighttime case, these results point to the error being introduced by the ICED model rather than the raytrace. This is because the interpolation scheme used in the raytrace is the same for both night and day.

## CONCLUSIONS

A year long campaign successfully produced a unique data base on the refractive bending of a LoVHF signal caused by a variety of ionospheric conditions. A ground-based direction finder used an orbiting beacon on 29.5 MHz to map ionospheric electron density. The effects of day/night cycles and season were observed. Short term variations, such as those due to sporadic E and ionospheric storms were measured.

During periods of electron density minimum, the vector sum of the azimuth and elevation errors ranged near 1.5 percent of the slant range between the satellite and receiver. This is likely the instrumental accuracy of the 7 channel interferometer that was used.

During measurements when the electron density was at maximum values, the transionospheric angular errors ranged between 5 and 10 percent of range. Short term gradients caused by sporadic E and ionospheric storms pushed the errors to excessive levels.

A relationship was developed that relates transionospheric angular error to the peak electron density as measured from a ground-based vertical sounder. This was then used to test various ionospheric models to see if they could first replicate the observed bending and then correct for it. The results were not satisfactory as the ICED model tends to over-predict the electron density during daylight hours. It is suspected that this is due, in part, to a poorly modeled topside ionosphere, from 350 km up to the 1000 kilometer altitude of the satellite.

## ACKNOWLEDGEMENT

The author acknowledges the contributions of Richard A. Sprague for innovative analytical work and Dr. David R. Lambert in the review and editing of this paper.

## REFERENCES

- Daniell Jr., R.E., D.T. Decker, D.N. Anderson, J.R. Jasperse, J.J. Sojka, R.W. Schunk, "A Global Ionospheric Conductivity and Electron Density (ICED) Model," Proceedings from the Ionospheric Effects Symposium, The Effect of the Ionosphere on Radiowave signals and System Performance, 1-3 May 1990
- Rose, R.B., "A Current Assessment of Single Site Locating Technology," Electromagnetic Wave Propagation Panel NATO-AGARD 50th Symposium on Radiolocation Techniques, London, UK, 1-5 June 1992.



## DISCUSSION

### T. COYNE

I understood from your previous presentation (paper 10) that problems above 82° elevation were associated with ionospheric tilts - I take it from your present paper this is not correct. The problem is an instrumental one - which surprises me as you are close to line-of-sight to the interferometer when it is equipped with loops.

### AUTHOR'S REPLY

For an interferometer, azimuth calculations become unstable at elevation angles of greater than 82°. This is systematic of the system that we are using. A fairly complex solution was developed by Frank Polkinghorn of NRL to circumvent the large azimuth errors at overhead angles.

### E. HAYDEN

For the interferometer system in question, when the signal arrives from directly overhead it is true that the azimuth measurement becomes indeterminate. However, in that same situation it is important to keep in mind that the elevation measurement is at its most accurate because the wave direction is normal to the array plane.

### AUTHOR'S REPLY

Concur.

### L. BERTEL

Calculation of angle errors requires knowledge of the electron density in the vicinity of the satellite: do you take that parameter into account in your simulations? And, if the answer is 'yes', what ionospheric profile do you use?

### AUTHOR'S REPLY

Initially we used rather simplistic assumptions in structuring the experiment. The beacon transmitter was at 1000 km altitude. At this altitude, the emitter is above 96-97% of the total electron content. Therefore, we initially assumed, the transmitter and receiver were completely above and below the ionosphere and the bending observed was due to the plasma in between. The ionospheric profiles we used were from ICED for the midlatitude experiment. With the type of day time errors observed, the median ionospheric electron content is too high already. To account for the residual 3-4% of electron content that exists between 1000 and 1500 km would only increase the error.

### C. GOUTELARD

Pourquoi n'utilisez-vous pas le modèle de Bent pour la modélisation de l'ionosphère au voisinage du satellite? (suite à la question de Mr Bertel).

*Why not use the Bent model for modeling the ionosphere in the vicinity of the satellite? (Follow-up to Mr. Bertel's question).*

### AUTHOR'S REPLY

Our project was specifically tasked to investigate useability of electron density profiles from ICED. We could have used the same process on any Ne profile from any model. It just wasn't our task on this project.

### H. SOICHER

Transionospheric transmissions at 29.5 MHz are occasionally subjected to anomalous propagation conditions (e.g., reflections to receivers from totally different directions). Have you noticed such? Under what circumstances?

### AUTHOR'S REPLY

Yes, myself and the scientists at SWRI, where the measurements were made, were very aware of what happens when signals at 29.5 MHz go skywave. One time on a north-south pass, we observed the beacon until it was over Antarctica. To the south, in the summer when the equatorial anomaly was the furthest north we saw many cases where large azimuth and elevation errors occurred. Normally, measurements to the north were well behaved. We have confirmed cases from ionospheric storms, E<sub>s</sub>, auroral backscatter, and equatorial E. However, because of the very extensive data base collected during normal conditions, anomalous propagation amounted to a very small percentage of the data.

# Systematic bearing errors of HF signals observed over a North-South propagation path

T.B. Jones<sup>1</sup>  
E.M. Warrington<sup>2</sup>  
J.E. Perry<sup>3</sup>

<sup>1</sup>Department of Physics and Astronomy

<sup>2</sup>Department of Engineering  
University of Leicester  
University Road  
Leicester, LE1 7RH  
UK

<sup>3</sup>Department of Defense, USA

## SUMMARY

Strong gradients in electron density associated with sunrise and sunset are a well known feature of the ionosphere and will produce off great circle bearings, particularly if the propagation path is parallel to the dawn or dusk terminator. Measurements have been undertaken of the direction of arrival of signals in the range 3 to 23 MHz radiated by a transmitter located at Clyde River in the Canadian Arctic (70°N, 70°W). The path from this transmitter to the receiving site at Boston, USA, is parallel to the dawn dusk line and consequently systematic changes in bearings are expected to occur. Bearings measured during January 1989 indicate a positive error of a few degrees at around sunrise. As the day progressed the error decreased, becoming zero at local noon at the path mid point. As the dusk approaches, the tilts in the ionosphere are reversed in gradient and there is a smaller negative swing in the mean bearing. The bearing error at dusk is smaller than at sunrise since the ionospheric gradients at this time are less steep. The diurnal swing in the bearing occurs during the winter and equinox periods but is absent (or very small) during summer. This is because the ionospheric gradients in summer are smaller than those at other seasons due to the relatively low values of the F-region critical frequency (foF2) which occur during the summer daytime.

## 1. INTRODUCTION

Errors in the angles of arrival of HF signals received after reflection from the ionosphere can occur because of gradients in the electron density height distribution. Such gradients produce a tilted reflecting surface which allows the signals to propagate from the transmitter to the receiver via off-great circle paths. Tilts and gradients are produced in the ionosphere by a number of geophysical processes, eg. the propagation of internal gravity waves which cause travelling ionospheric disturbances. This investigation is concerned with the influence of the large, slowly moving, tilts which are associated with the dawn/dusk terminators. These tilts are found to produce systematic errors in bearings, particularly for paths whose direction is parallel to the line of the terminator.

Bearing errors of 2° to 3° resulting from the sunrise/sunset gradients in the ionosphere have been observed by Tedd et al (Reference 1) for a number of one-hop paths in Western Europe. In this study, the transmission path is located in the North American sector.

## 2. EXPERIMENTAL RESULTS

An HF transmitter has been located at Clyde River, North West Territories, Canada, and the bearings of the signals measured at a site near Boston, Massachusetts, in the U.S.A. The frequencies employed ranged from 3 MHz to 23 MHz and the results reported here are for frequencies greater than 10 MHz. The upper frequencies have been selected, since they were above the LUF and below the MUF during most of the dawn/dusk periods. Bearing measurements were made using a wide aperture goniometric direction finder and the observed azimuthal angle of arrival was recorded at 30 second intervals during each 2 minute transmission interval. The transmissions were repeated at half-hour intervals during the observing periods which covered 30 days during the summer, winter and equinox periods.

The results obtained during January 1989 are reproduced in Figure 1 in which the average bearing for each hourly interval for all frequencies above 10 MHz has been plotted. The signals are first acquired at sunrise when the path opens due to the increasing electron density caused by the solar illumination. This results in a rapid increase in MUF to a value well above 10 MHz. At this time, there is a positive error of about 2° in the bearing measurement. As the day progresses, the error decreases, becoming zero at local noon at the path mid point (17:20 UT). At this time, the ionosphere has fully formed and there are no large E-W gradients present. As dusk approaches, the electron density decreases due to the increasing domination of recombination. The E-W tilt is, therefore, in the opposite direction to that at sunrise. This can be seen as a small 1° negative error in the bearing measurements. The smaller error at dusk is to be expected since the gradients at this time are less steep than at sunrise. This can be seen from Figures 2(a) and (b), in which the

contours of constant electron density (iso-ionic contours) are reproduced for summer and winter conditions respectively. At latitudes between 45° and 70°N, the sunrise gradient in winter is much steeper than at sunset. In summer, the difference is not so marked because of the lower values of the F-region critical frequency which occur in summer daytime. The transition from night to day is, therefore, less marked during this season. This conclusion is confirmed experimentally and the systematic bearing error during summer is too small to measure.

### 3. DISCUSSION

The bearing error due to a tilted ionosphere can be determined by ray tracing analysis. Ray tracing calculations for the Clyde-Boston path have been undertaken by means of the Jones (Reference 2) analysis. The model ionosphere assumed in these calculations have been based on the data contained in Figures 2(a) and (b). Positive bearing errors of about 2 or 3 degrees are obtained from the winter model (Figure 2(a)) whereas very small errors (less than 1 degree) are produced by the gradients present in the summer model (Figure 2(b)). These calculations confirm the experimental evidence that systematic bearing errors are associated with the sunrise/sunset gradient particularly in winter.

The present results and those of Tedd (Reference 1) suggest that these bearing errors are a regular feature associated with sunrise and sunset. Thus, a simple correction (addition and subtraction) could be applied to the measured bearings for this type of propagation path. Note that the maximum error will occur for North - South paths since the electron density

gradient is then orthogonal to the path direction. As the path direction moves away from N - S, the bearing error due to these tilts becomes smaller and will become zero for E-W propagation when this path direction is parallel to that of the tilt. The magnitude of these errors is not critically dependent on path length provided the transmission is via a one-hop mode.

### 4. CONCLUSIONS

The gradients in electron density which are produced in the ionosphere at sunrise and sunset can influence the angle of arrival of HF signals. This is particularly true for paths which are perpendicular to the gradient since maximum deflection of the reflection point from its plane mirror position will occur when the tilt is at right angles to the path. The gradient, and hence the bearing error, are greater at sunrise than at sunset and have their greatest magnitude in winter. In summer, when the difference in the magnitude of the night-time and daytime electron density is smaller, little change is observed in the bearing. The systematic nature of these errors and their repeatability from day to day suggests that a correction procedure could be developed for their removal.

### 5. REFERENCES

1. Tedd B L, H J Strangeway and T B Jones, Systematic ionospheric electron density tilts at mid latitudes and their associated bearing errors, *JATP*, 47, 1085, 1985.
2. Jones R M, *A 3-D Ray Tracing program*, ESA Technical Report 1ER17, ITSA17, 1966.

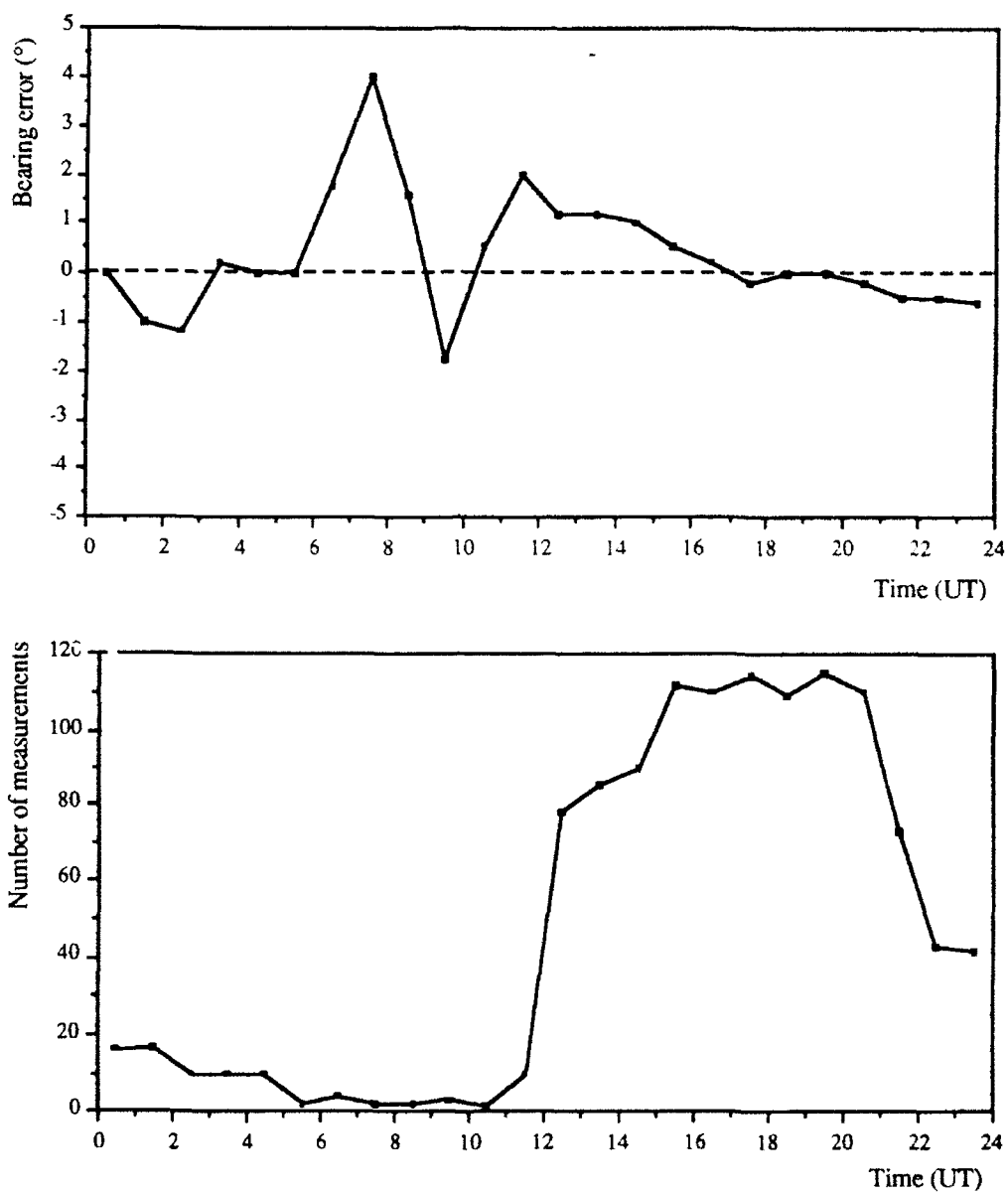


Figure 1. DF measurements for the Clyde River - Boston path. January 1989. The upper frame shows the average bearing error as a function of time of day for all frequencies above 10 MHz and the lower frame the number of bearing measurements made.

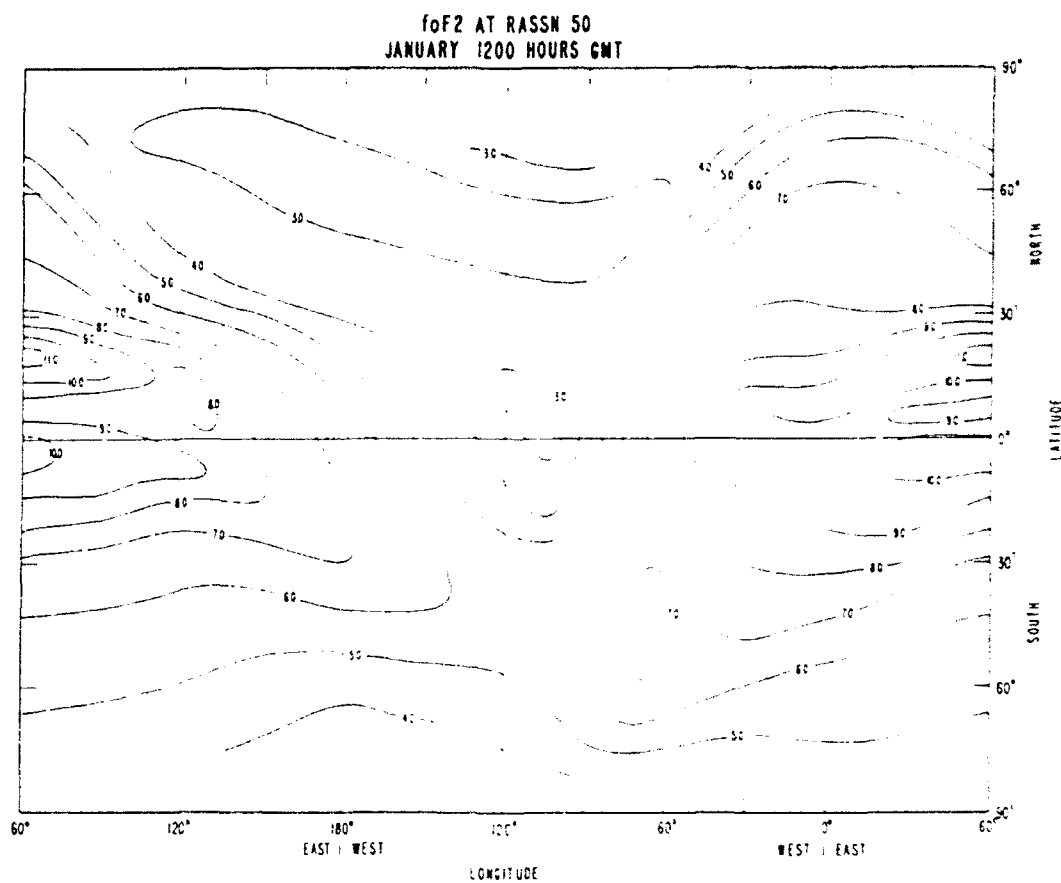


Figure 2 (a). World maps of foF2 for winter. Note the steep East-West gradients, especially at sunrise.

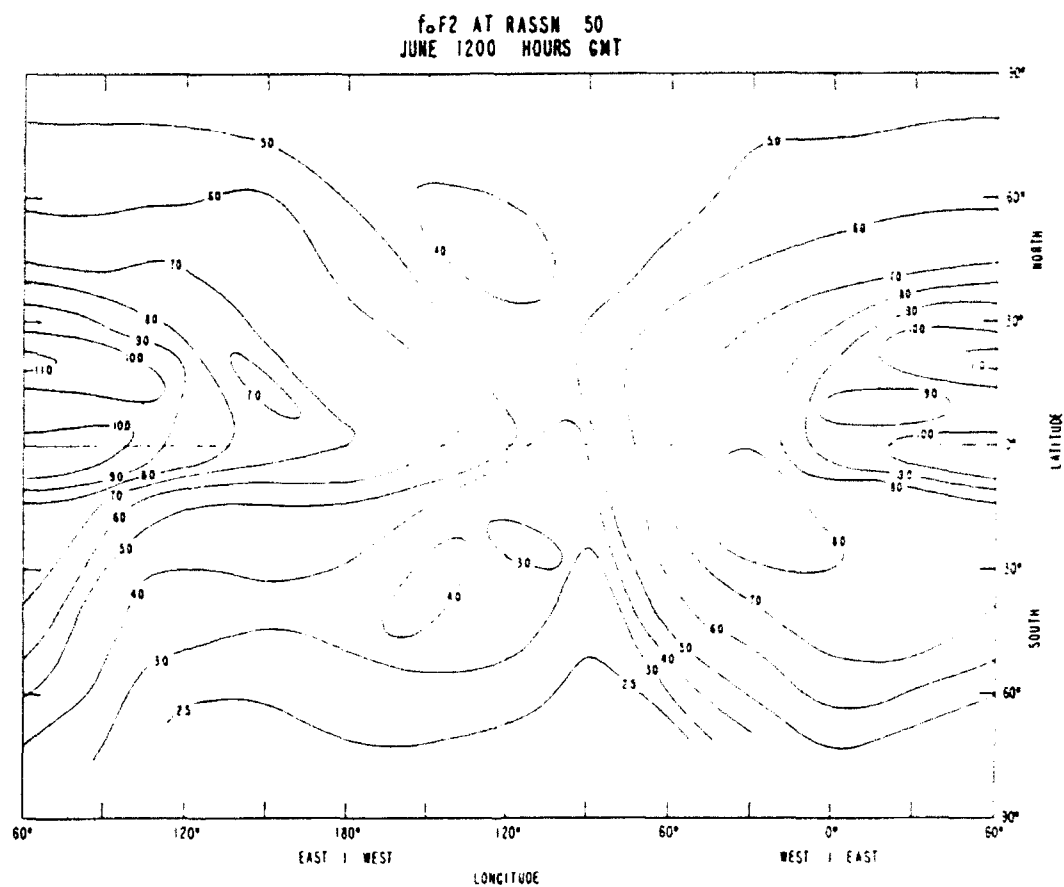


Figure 2 (b). World maps of  $f_oF_2$  for summer. Note the small East-West gradients at sunrise and sunset.

# Large bearing errors observed at a high latitude DF site

T.B. Jones<sup>1</sup>  
E.M. Warrington<sup>2</sup>

<sup>1</sup>Department of Physics and Astronomy

<sup>2</sup>Department of Engineering  
University of Leicester  
University Road  
Leicester, LE1 7RH  
UK

## SUMMARY

It is well known that the auroral zone ionosphere contains steep time varying gradients in electron density which can produce large off-great circle bearings in the direction of arrival of HF radio signals reflected from this region of the ionosphere. Recent DF measurements at a high latitude site indicate that large bearing errors are also observed in signals reflected from the polar cap ionosphere which is generally regarded as being a much less disturbed region than the auroral zone. A feature of these measurements is a large ( $\pm 50^\circ$ ) quasi-periodic swing in the direction of arrival observed for both short and long propagation paths. These changes in angle of arrival are attributed to reflection from patches or "blobs" of over-dense plasma which travel with the convection current flow across the polar cap region from dayside to nightside. The periods of large bearing swings are well correlated with the onset of magnetic activity as measured by the Kp index.

## 1. INTRODUCTION

The errors measured by HF direction finders on signals received over high latitude propagation paths are usually much greater than those of signals received via the mid and low latitude ionosphere. These errors have generally been attributed to the propagation path reflection points being in or near the auroral oval where large gradients in electron density are known to occur. These gradients form tilted reflection surfaces which allow off great circle propagation paths between the transmitter and the DF site. In November 1990 a controlled series of experiments were undertaken in the high Arctic, at Alert, North West Territories, Canada, (82.5°N, 62.4°W) with a goniometer based and with an interferometer based DF system. An interesting feature of the data collected during the tests was that on some days very large bearing swings (in excess of  $\pm 50^\circ$ ) were measured by both instruments and on two propagation paths. Furthermore, the interferometer also indicated changes in the elevation angle of arrival although these were not, in general, well correlated with the bearing changes. These large errors were measured on the Thule-Alert path (676 km) which is always contained entirely within the polar cap and on the Halifax-Alert path (4200 km) which has one of its reflection points within the polar cap.

Since the polar cap ionosphere is known to be less disturbed than that in the auroral oval and, moreover, there is no obvious source of ionospheric gradients within the polar cap region, a more detailed investigation of these observations was undertaken with a view to identifying the cause of the bearing errors. A possible cause of these errors is identified and the geophysical parameters responsible for the disturbance discussed.

## 2. EXPERIMENTAL OBSERVATIONS

During the period 11-20 November 1990, measurements were made at Alert of the arrival angles of the signals received from Thule (8.050 MHz) and Halifax (8.697 MHz). These transmitters are located at 676 and 4200 km from Alert respectively (see Figure 1).

Both the azimuth data collected by the goniometer system and the azimuth and elevation angles measured by the interferometer indicate that large time varying changes in these parameters are frequently observed. Typical examples of the large azimuth changes recorded by the goniometer system on 18 November 1990, are reproduced in Figure 2 (a) and (b). In Figure 2 (a) the bearings of the 8.050 MHz Thule transmission are plotted as a function of the time of day (UT). Similar azimuth angle data measured by the same system for the Halifax (8.697 MHz) signal are reproduced in Figure 2(b). Very large bearing swings, of more than 100 degrees, are evident on both paths during the period 08.00 to 12.00 UT (ie. 04.00 to 08.00 local). Note that the swing is from a negative bearing error through the true great circle position to a positive error in both cases. The true bearings for the Halifax and Thule transmitters are 182 and 194 degrees respectively and these are marked on the figure. The period between 12.00 and 16.00 UT is relatively undisturbed but after 18.00 UT (14.00 local), large bearing errors are again observed on the Thule path. At this time, however, the bearing moves from a large positive value, through the great circle position, to a large negative value, as time progresses. Note that the number of bearing swings is fewer in the post noon than in the pre noon periods and that large bearing errors are observed only on the Thule path in the post noon period.

The azimuth and elevation angle data measured by the interferometer during the 17 November are reproduced in

Figure 3 (a and b) for the 8.050 MHz transmission from Thule. High time resolution plots of these data indicate that two or more bearing angles can occur simultaneously but the dominant feature is the swing from small to large bearings as time progresses. The changes in the elevation angle are less well defined than those in azimuth and there is a considerable ( $\sim 15^\circ$ ) spread in this parameter.

### 3. INTERPRETATION OF THE DATA

#### 3.1 Production mechanisms for large bearing errors

The large periodic swings in azimuth, for example on the 17 and 18 November, indicate that the the incoming wave direction varies from approximately  $50^\circ$  on one side of the true bearing to  $50^\circ$  on the other side of true bearing within about 30 mins. The azimuth change is accompanied by changes in elevation angle which do not appear to be directly correlated with the bearing change.

To produce off great circle path propagation, tilts and/or gradients in electron density are required. If the bearing changes with time, then either the gradient must change or the reflection point must move relative to the transmitter-receiver direction. Gradient changes associated with the auroral oval or the high latitude trough can give rise to large bearing errors, however, these gradients are slow moving and cannot produce the quasi periodicity which are a feature of the present observations. The most common disturbances in the ionosphere, which gives rise to time variations similar to those observed, are those associated with travelling ionospheric disturbances (TIDs). TIDs can occur at any latitude and at any time of day or season. The period of these disturbances ranges from 20 min to more than 1 hour. However, the electron density gradients associated with TIDs are relatively small and will typically produce off great circle bearings of about  $\pm 5^\circ$  for a one hop propagation path (Reference 1). Even assuming that TIDs in the auroral oval were larger than normal, they could not produce the very large  $\pm 50^\circ$  bearing errors observed.

Gradient features specifically related with the polar cap region are those associated with convecting plasma 'blobs' or patches of enhanced electron density. Recent observations (Reference 2) of the polar cap ionosphere have established the existence of large blobs of over-dense plasma which convect across the polar cap region carried along by the convection current flows. These flows usually form a two-cell convection pattern with the flow over the pole being from the dayside to nightside, (see Figure 4). On the dayside, during moderate geomagnetic activity, parts of the enhanced auroral oval plasma can become detached. Due to the long recombination time-constants in the F-region this detached plasma will have a long life time (several hours) and so will move across the polar cap embedded into the normal convection flow. The feature will then appear as a patch or 'blob' of over-dense plasma moving through the ambient polar cap ionosphere.

These features have been observed by high latitude incoherent scatter radars such as Sonderstromfjord, Millstone Hill and EISCAT. The electron density within the blobs is considerably greater than that of the ambient ionosphere. Consequently, the blob boundaries correspond to regions of steep gradients in electron density which act as good tilted

reflectors for HF radiowaves. It should be noted that the over-dense plasma regions can extend laterally for several hundred kilometres and extend in height throughout the F-region. The features of individual blobs can vary considerably and their density will gradually decrease as recombination taken place.

The physical processes which produce the detached plasma patches are not well understood but are known to be associated with the dayside cusp region of the magnetosphere, with the level of geomagnetic activity and the magnitude and direction of the interplanetary magnetic field (IMF). In order to account for the DF observations it now has to be established that time varying bearing errors can be produced by the convecting blobs.

#### 3.2 Modelling of the bearing changes due to a plasma blob

Let the blob be of sufficiently high electron density to reflect the radiowave transmitted from T and received at R - see Figure 5. Assuming that the electron density gradients in the blob are such that they are capable of reflecting the signal, a propagation path from T to R via the blob will be established as soon as the blob is illuminated by the transmitter (position A). This will produce a bearing error  $\theta_1$ . As the blob moves along the direction AB, the bearing error will decrease and the elevation angle increases. The signal will be received along the great circle position when the blob is at point D. The bearing will then increase in the opposite sense ( $\theta_2$ ) as the blob travels. The maximum elevation angle will be recorded at the point of closest approach to the receiver, ie. at C. The bearing angle error continues to increase as the blob moves to the radio horizon of the transmitter at B. This sequence depends on the blob always presenting a suitable gradient for propagation to occur between the transmitter and receiver. This would require convex reflecting surfaces, as would be the case in an over-dense blob.

Since the exact form of the blob is not known, and will, in any case, vary from blob to blob, a somewhat simplified approach has been adopted in order to model the variation of the azimuth and elevation angles with time during the transit of the blob. It has been assumed that the blob is a point specular reflector which moves along any arbitrary linear path. The height of the reflector can be varied as can the path direction and the velocity of transit. The computer simulation indicates that changing the height of the reflector (100 and 300 km) makes little difference to the form the variation in either parameter. The magnitude of the changes are, however, greater when the higher (300 km) reflection height is adopted.

As might be expected, the position of the reflector relative to the receiver at its point of closest approach, makes a considerable difference to the rate of change, particularly in the bearing angle. Furthermore, if the reflector were to pass directly over the receiver, any azimuthal angle of arrival could be generated. The above modelling serves only to illustrate the kind of changes that may be expected and such changes have been observed in the measured data. For a real blob, neither the reflection height nor the reflection point on the blob would remain constant as the blob travels and these changes would influence both the azimuth and elevation angles measured at the receiver. The changes in reflection geometry between transmitter and receiver would also have to be taken into account as would the size of the blob (several



hundred km). For example, the blob may not travel along a straight line path and there may be more than one blob feature present at any given time. A more realistic model of this kind is now being developed.

### 3.3 Diurnal changes

One of the interesting features of the large bearing swings observed on both 17 and 18 November on the Thule-Alert path is that the direction of rotation of the bearing error is different in the pre and post noon periods. Around local noon there is little change in bearing which is near to its great circle value. Thus, the bearing changes from small to large values in the morning and from large to small values in the afternoon. This implies that the direction in which the reflecting blob crosses the great circle path between the transmitter and receiver must change between the two periods.

A basic model of the polar cap convection flow is reproduced in Figure 4. The two-cell structure is clearly evident as is the anti-sunward flow over the magnetic pole. This pattern is fixed in space (relative to the sun) and the earth rotates beneath these flows. For the Halifax and Thule to Alert paths, local and magnetic times are quite similar, thus, in the morning the flow is almost orthogonal to the path (T - R) as indicated in Figure 6 (a). Blobs detached from the dayside (~12.00 UT) will follow the direction of the convection flow over the polar cap as indicated by the arrow. Reflections from the blob will thus be received first when it is in position (1), i.e. the measured bearing will be smaller than the great circle value. As the blob moves with the flow, the bearing angle will increase to its true great circle (zero error) position and continue to increase as the blob travels away to position (2).

Near noon, Figure 6 (b), the flow is approximately parallel to the propagation path, consequently the presence of blobs will have little effect on the azimuthal angle of arrival although they will still produce changes in the elevation angle.

In the post noon period, the flow is again almost perpendicular to the path but this time the earth has rotated so that the path is in the position shown in Figure 6 (c). When reflections from the blob are first obtained, position (1), the bearings will be larger than the great circle values. These values will decrease as the blob moves with the convection flow, the smallest bearing angles being recorded at position (2), when the path from the blob finally disappears.

The change in position of the propagation path relative to the Ionosphere/Magnetosphere convection flow as the earth rotates, can account for the change in the direction of rotation of the measured bearing errors between the pre and post noon periods. The number of bearing swings will depend on the number of blobs present that can provide reflection points between the transmitter and receiver. The period of the bearing swing will depend on the velocity of the blob relative to the transmission path. This will depend on several factors including the velocity of the flow vectors which, in turn, depend on the magnitude and direction of the IMF and the cross-cap potential.

### 3.4 Correlation with other geophysical observations

(i) The Thule ionosonde. The presence of large blobs of over-dense plasma should register on a vertical incidence ionosonde located in the convection flow region as a

substantial increase in the F-region critical frequency while the blob is within the first Fresnel zone of the ionosonde. The duration of the F-region critical frequencies enhancement will, of course, depend on the velocity of the blob. The time plot of foF2 should, therefore, contain quasi periodic enhancements during blob activity. Such plots are reproduced in Figure 7a and 7b from data produced by the Thule ionosonde (17 and 18 November 1990) (Reference 3). For contrast, the variation of foF2 for a quiet day (19 November 1990) is presented in Figure 7c.

There is a very clear difference between the F-region behaviour as measured by the ionosonde on the blob active and quiet days. The data presented in Figure 7 confirm that large transient enhancements in electron density occurred on 17 and 18 November whereas, in contrast, no such changes were present on the quiet day (19 November). Comparison with the DF data indicates that large bearing errors are only present on those days when the ionosonde records large transient fluctuations in foF2. Good bearing performance is always obtained during quiet days when there are no disturbances in the ionosonde records. Detailed comparison between the measured foF2 and bearing fluctuations indicate that there is not a high degree of correlation between individual events. This may be due to the location of the ionosonde not being at the propagation path mid point or that the blobs might change their form as they propagate from the location of the ionosonde to that of the DF path reflection point.

The study of the ionosonde data revealed that the disturbances in critical frequency occurred only during periods of relatively high geomagnetic activity. The 3 hour Kp values have been included in Figure 7 and it is evident that these are much higher for the active days compared with the quiet day. It appears that an Ap value of 10 or greater can lead to the type of fluctuation of foF2 associated with propagating blobs. This suggests that the magnetometer data might provide an indication of blob activity and, hence, of the likely occurrence of large bearing errors.

(ii) Magnetometer measurements. It has already been established that days of moderate magnetic activity ( $A_p > 10$ ,  $K_p > 3$ ) were likely to have blob-like features which could be detected by the Thule ionosonde in foF2 and by the DF systems at Alert as large bearing errors on the polar cap propagation paths. To enable closer inspection of the magnetic disturbances, magnetometer measurements were obtained from four sites in Northern Canada (St Johns, Great Whale, Resolute and Alert) which covered a range of geomagnetic latitude from 60° to 86°N.

The large bearing errors were observed on days 320 and 321 (17 and 18 November) in the periods 8 to 12 and 18 to 24 UT. The 17 November (day 320) corresponds to the onset of magnetic activity at all four magnetometer sites and this is particularly evident in the Great Whale data. At Resolute and Alert these disturbances occur in the afternoon whereas the bearing errors were present during both pre and post noon periods. The disturbances at Great Whale, however, occur in the morning period some 2 or 3 hours before the large bearing errors were recorded at Alert. The Great Whale magnetometer is located under the auroral oval at this time and is, therefore, sensitive to current flows (conductivity

changes) within the oval. These disturbances could, therefore, be associated with changes in the oval which lead to plasma break-up on the dayside and, hence, to blob production. Further investigation is needed of the relationship between the magnetometer disturbances, the production of blobs and the incidence of large bearing errors in DF measurements at Alert.

#### 4. SUMMARY AND CONCLUSIONS

During DF measurements at Alert, large quasi ( $>\pm 50^\circ$ ) periodic bearing changes were observed on a number of occasions. These large errors were observed on a short (676 km) path contained entirely within the polar cap and on a longer (4200 km) trans auroral path which had a reflection point in the polar cap. For the short path, the direction of change of the bearing error was different in the pre and post noon periods.

Possible causes of these large errors have been examined and evidence is presented which suggests that they are due to the convection of large blobs of over-dense plasma across the polar cap. Modelling of the moving reflection point yields changes in the azimuth and elevation angles similar to those observed experimentally. Moreover, the blob movement along the convection flow would account for the difference in

direction of the bearing swings in the morning and afternoon periods. The presence of 'blobs' has been related to disturbances on (a) the vertical incidence ionosonde at Thule and (b) a number of magnetometers located in Northern Canada.

#### 5. ACKNOWLEDGEMENT

The authors would like to thank the Canadian Department of National Defence for their support of this work.

#### 6. REFERENCES

1. Tedd B L, H J Strangeways, and T B Jones The influence of large-scale TIDs on the bearings of geographically spaced HF transmissions. JATP, 46, No.2, p. 109 - 117, 1984.
2. Weber E J, Auroral zone plasma enhancements. JGR, 90, 6497-6573, 1985.
3. Buchau J, Ionosonde data (Private communication), 1991.

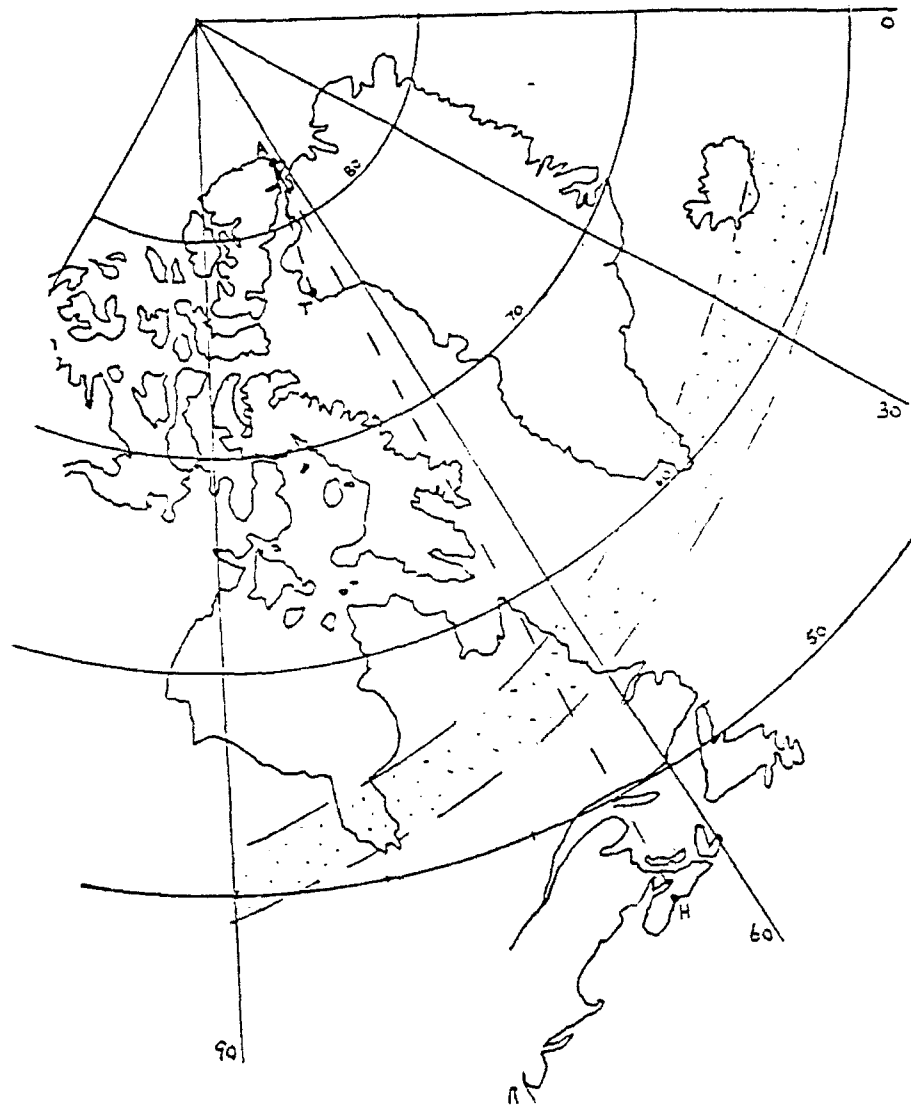


Figure 1. Location of the DF (Alert) and transmitter (Halifax and Thule) sites relative to the auroral oval.

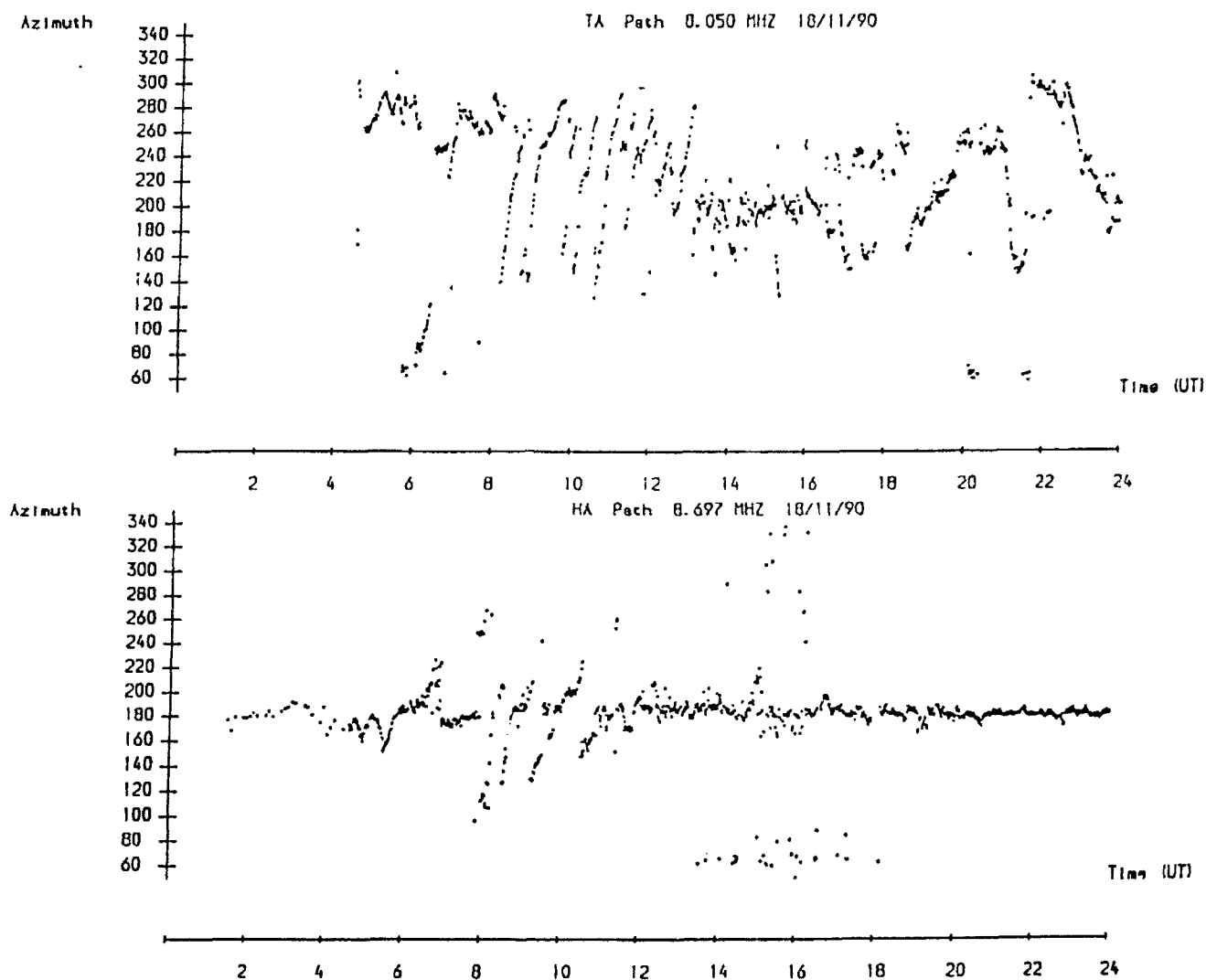


Figure 2. Large bearing errors observed on the Thule - Alert (TA) and Halifax - Alert (HA) paths, 18 November 1990.

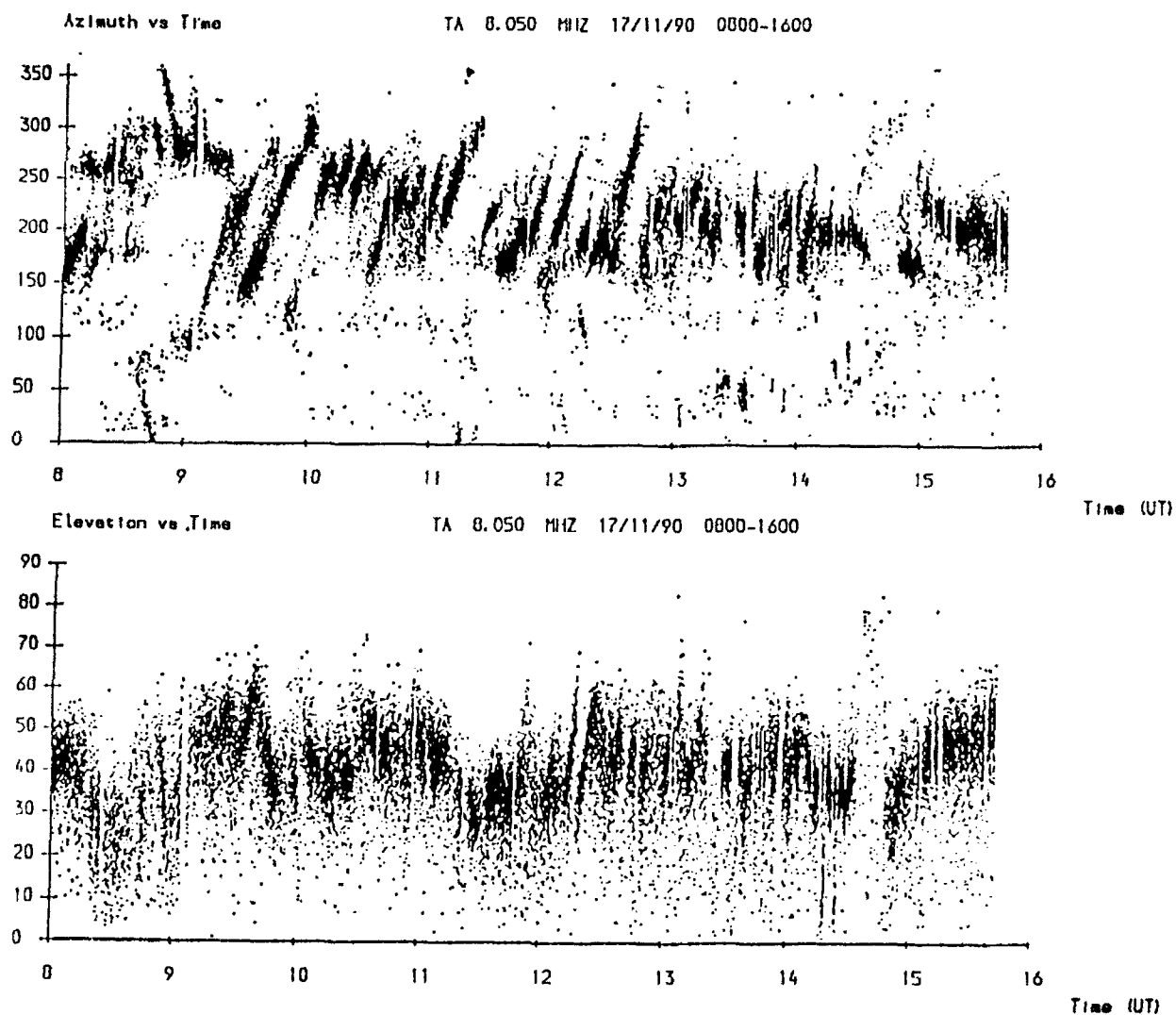
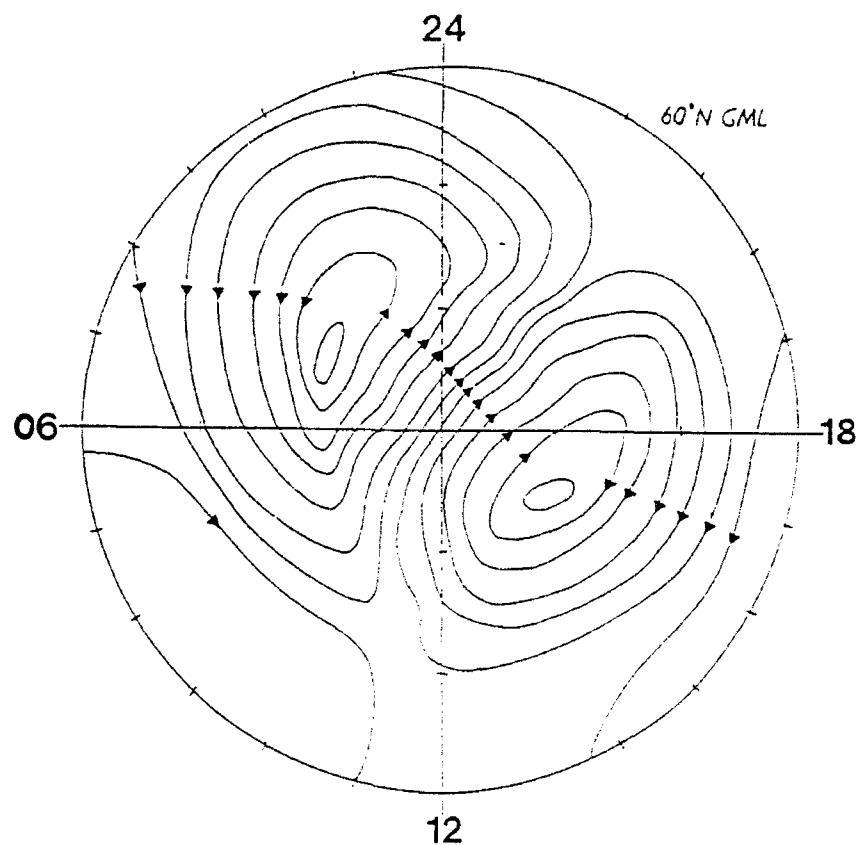


Figure 3. Interferometer measurements of the (a) azimuthal and (b) elevation angles of arrival of the 8.050 MHz signal from Thule, 08.00 to 16.00 UT, 17 November 1990



PLASMA CONVECTION PATTERN FOR  $B_z$  NEGATIVE  
(FROM FRIIS-CHRISTENSEN ET AL 1985)

Figure 4. Plasma convection flow over the polar cap showing the two cell pattern with dawn to dusk flow over the magnetic pole.

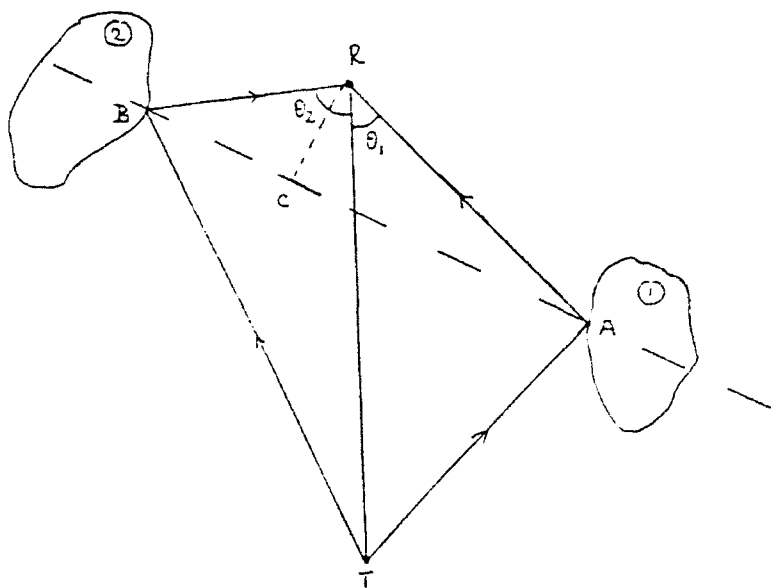


Figure 5. Schematic diagram of plasma 'blob' movement relative to the great circle propagation path (T to R). Note the change in the off great circle path as the 'blob' convects.

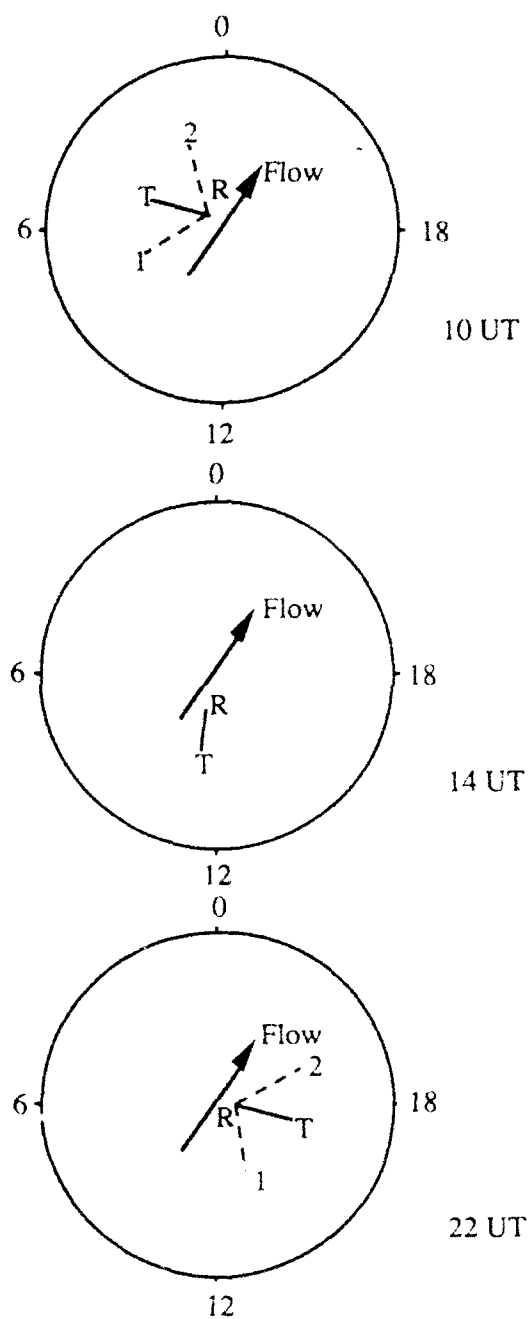


Figure 6. Schematic diagram of the great circle path (TR) relative to the mean convection flow direction for pre-noon, noon and post-noon periods.

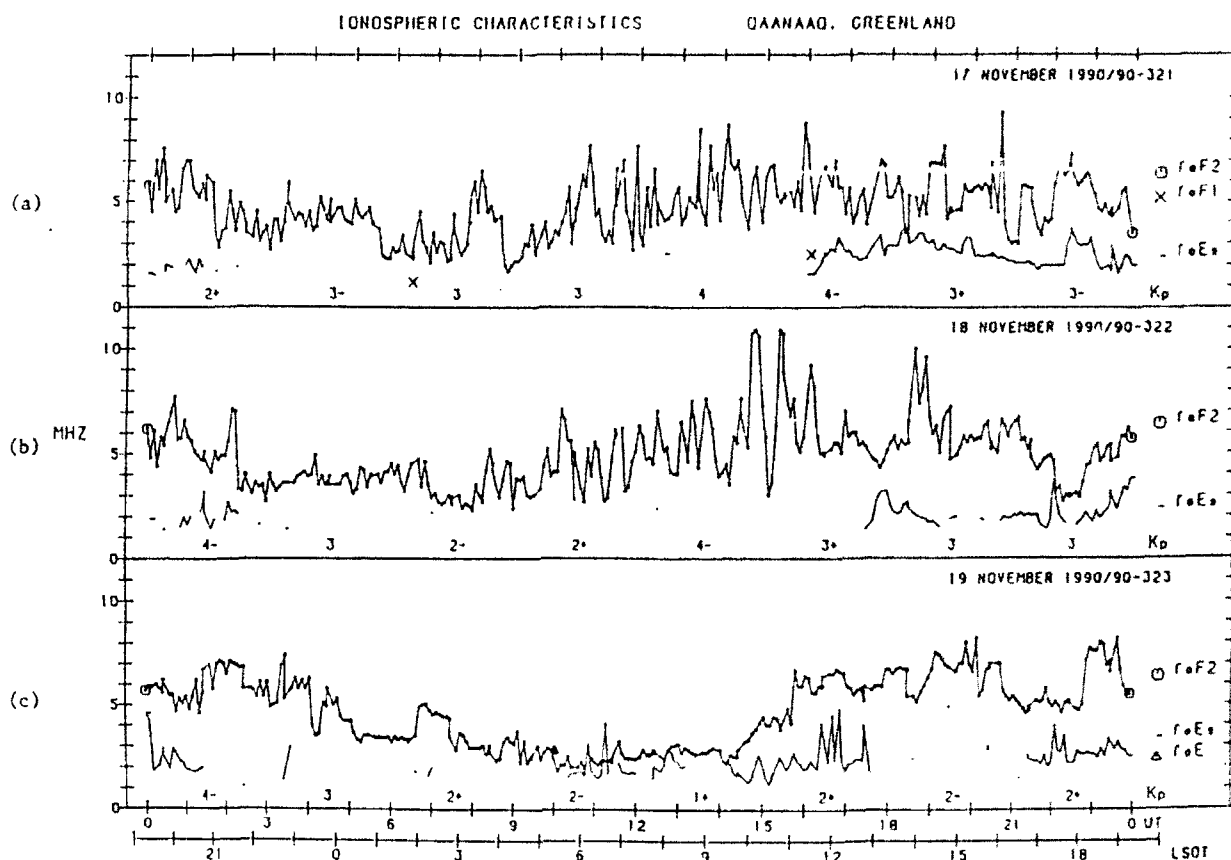


Figure 7. Changes in F-region critical frequency observed at Thule during periods of blob activity (17, 18 November 1990) when large bearing errors were measured at Alert. An undisturbed day (19 November 1990) is included for comparison.

## DISCUSSION

### R.W. JENKINS

#### Comments:

1. At the end of your presentation, you mentioned that on the east-west path, between Halifax and Ottawa, the observed bearing swung to the north for several hours in local evening on the days when the large swings are observed at Alert. This could be due to a reflection from the northern edge of the ionospheric trough, coupled with a drop in the MUF for the great-circle path causing that path to be absent for that time. Such an effect was observed by us using the Ottawa HFDF array, on signals from an aircraft flying over the North Atlantic region near Newfoundland, and is reported in a previous AGARD-EPP conference ('Direction and Doppler Characteristics of Medium and Long Path HF Signals Within the Night-Time Sub Auroral Region,' R.W. Jenkins, E.L. Hagg, and L.E. Montbriond, Paper 20, AGARD-CP-233, 1979).

2. We have looked at the same high-latitude data taken by interferometer, and also interpret the observed motions to be due to convecting patches. However, we expect that there will be a great many reflecting points or scattering points within these patches, due to turbulences caused by the gradient-drift instability - which should cause the spreading in both bearing and elevation between consecutive interferometer measurements. We were able to infer the locations of the convecting patches, as well as motions, between 300 and 800 m/sec, and the directions of those motions to be in agreement with the Friis-Anderson 2-cell convection patterns.

### D. HAINES

Did you ever see a great circle bearing angle superimposed upon the deviated path or is the bearing angle usually dominated by the off-great-circle path when plasma patches exist?

### AUTHOR'S REPLY

In 10 days of data we observed no case in which an off-great-circle and great-circle-path were superimposed. When present, the signals propagated via the plasma patches did dominate the bearing angle.



# Measurements of the direction of arrival of an oblique chirp sounder signal

E.M. Warrington<sup>1</sup>

T.B. Jones<sup>2</sup>

P. Hamadyk<sup>2</sup>

<sup>1</sup>Department of Engineering

<sup>2</sup>Department of Physics and Astronomy

University of Leicester

University Road

Leicester, LE1 7RH

UK

## SUMMARY

The performance of high frequency (HF) direction finding systems is related to the mode content of the received signal and to the signal frequency. In order to fully investigate this effect, a novel experiment has been devised to measure the direction of arrival of oblique ionospheric sounding signals emitted by the worldwide network of BR Communications chirp sounders. These signals are radiated as a constant sweep chirp signal from 2 to 16 or 30 MHz thus enabling the DF performance to be measured over the full range of frequencies propagating from the chosen transmitter. Simultaneous oblique ionograms are also recorded for the paths of interest in order to determine the mode structure at any frequency.

Preliminary tests of the system have been undertaken at Ottawa, Canada and Cheltenham, UK with the directions of arrival measured by a wide aperture goniometric DF system. Several interesting features have been identified in these measurements. A further measurement campaign has been conducted during April 1992 at Alert, a very high latitude site in the Canadian Arctic where very large bearing errors and systematic bearing swings, sometimes in excess of  $\pm 50^\circ$ , are known to occur.

## 1. INTRODUCTION

The performance of high frequency direction finding systems is known to be related to the mode content of the incoming signal and to the frequency of operation. Furthermore, signals propagating via the high latitude ionosphere are known to be particularly disturbed with, for example, the signal often arriving at the receiver over paths well displaced from the great circle (see Reference 1 reproduced elsewhere in these proceedings). Observations at Alert, located high in the Canadian Arctic, have indicated that large off great circle bearings on fixed frequency signals are associated with definable features on oblique ionograms taken over the paths of interest. In order to further investigate these phenomena, a system has been developed which can measure the bearings of signals radiated by the worldwide network of BR chirp

sounder transmitters across the full range of propagating frequencies.

Preliminary tests of the system were undertaken in Ottawa during August 1991 and at Cheltenham during January 1992 with further measurements at Alert during April 1992. Results of a preliminary nature from the January 1992 test are presented here. Further experiments and analysis are required to fully understand the influence of the mode structure on the measured DF bearing accuracy.

## 2. EXPERIMENTAL ARRANGEMENT

### 2.1 BR Communications chirp sounders

The RCS5A is a sweeping HF receiver which tunes through the HF spectrum in synchronism with a transmitter at the far end of the circuit under investigation. The transmitter emits a CW signal which starts at 2 MHz and sweeps upward in frequency at a constant rate of 50 or 100 kHz per second for 280 seconds finishing at either 16 or 30 MHz depending upon the sweep rate. An internal clock in the RCS5A receiver starts its sweep synchronously with the transmitter and precisely tracks the sweeping transmitter signal.

The radio signal propagates by various modes (eg. 1 hop, 2 hop, etc.) which have different propagation delays. These delays correspond to the time required for the signal to travel from the transmitter to the ionosphere and back to the receiver. Assuming that the transmit and receive sweeps are started at exactly the same time, the RCS5A receiver sweep will have advanced to a slightly higher frequency by the time that the transmitted signal arrives at the receiver. This received frequency difference is converted to an audio baseband tone in the receiver. Tones of increasing frequency indicate increasing time delay of the ionospherically propagated signal. In practice, multiple tones are present in the sounder's baseband audio signal which represent the various delays caused by the different propagation modes. These multiple tones are processed within the RCS5A into discrete components which correspond to the various ionospheric modes present in the received signal. The processed signal spectra are displayed as an oblique ionogram

on the RCSSA front panel display. In addition, these data are output on a serial RS232 link, thus enabling the ionograms to be collected and stored with the aid of a computer system.

Further details of the BR chirp sounder system may be found in Reference 2, together with details of ROSE, an enhanced signal processing facility which is to be employed on future measurement campaigns.

## 2.2 DF6 direction finding system

The DF6 is an automatic HF direction finding system capable of operation in the frequency range 3 - 30 MHz. The antenna array is a Plessey PUSHER array comprising two concentric 24 element circular arrays with diameters of 75 m and 25 m. The larger of the two arrays (the lowband array) operates at frequencies of between 3 MHz and (usually) 10 MHz, and the smaller of the arrays (the highband array) at frequencies between (usually) 10 MHz and 30 MHz. At Ottawa and Alert, both of the arrays consist of single element vertical monopoles whereas at Cheltenham the low band array comprises doublet antennas to give an improved beam pattern.

RF signals from the sum beam output of a PUSHER antenna are demodulated in the system receiver and combined with goniometer positional information for initial processing. An average beam pattern from successive rotations of the goniometer is then formed. These average patterns are known as composite scans. Bearings are then computed by applying software algorithms to the digitised composite goniometer sum beam patterns. Further details of the signal processing can be found elsewhere in these proceedings (Reference 3).

## 2.3 Sweeping DF configuration

A block diagram of the configuration of the Sweeping DF system is presented as Figure 1. The BR sounder receiver is connected to an omni-directional antenna of the type employed for the elements of the DF array (elevated feed vertical monopole). This type of antenna was chosen to ensure that a close match is achieved between the elevation angle sensitivity of the antennas from which the bearing of the received signal and the antenna employed for the mode content estimation. The oblique ionograms are output from the RCSSA in digital form and stored on a PC for subsequent analysis. Future experiments will employ the ROSE signal processing unit (Reference 2) to obtain higher resolution ionograms from the RCSSA.

The local oscillator signal from the chirp sounder receiver is fed into a signal frequency converter unit which converts the variable frequency chirp sounder signal being received on the DF array into a constant frequency signal of 29.750 MHz which is input to the DF system. A second PC is used to task the DF to provide composite scan information (0.5 second averages) which are returned to the PC for storage and subsequent analysis. A signal is output from the RCSSA at the start of each sweep which is sensed by the DF tasking PC to provide synchronisation between the two parts of the system.

## 3. EXPERIMENTAL OBSERVATIONS

At the time of writing, the Sweeping DF system has been deployed at Ottawa, Canada and at Cheltenham, UK and several days test of data collected. Examples of these

measurements for the path from Oslo to Cheltenham are now presented.

### 3.1 Ionograms and bearing measurements

Figures 2, 3 and 4 show examples of the oblique ionograms and associated bearing measurements (0.5 second integration time) for three sweeps received over the 1200 km path from Oslo to Cheltenham at 02:30 UT on 12 January 1992, 11:00 UT on 9 January 1992 and at 12:00 UT on 8 January 1992 respectively. The upper frame of these figures display the measured bearing of the chirpsounder signal with a 0.5 second integration time and in the lower frame are the corresponding ionogram measurements. The y-axis scale indicating the path delay has an arbitrary zero since it is not possible to perfectly synchronise the chirpsounder receiver with the transmitter. Each of the ionogram traces is coded on a grey scale corresponding to the received signal amplitude for each mode. The true bearing of the Oslo transmitter is approximately 31°.

The ionogram of Figure 2 shows predominantly sporadic E (Es) propagation over the full HF band. A spread F region trace is apparent up to a frequency of around 8 MHz. A large bearing spread of about  $\pm 2^\circ$  is evident from around 9 MHz to 30 MHz with some evidence of a frequency dependence corresponding to the sporadic E mode. Between 6 and 9 MHz very few bearings are returned by the DF.

Figure 3 illustrates a period of predominantly multiple hop F-region propagation with an MUF of about 26 MHz. Large bearing spreads of  $\pm 5^\circ$  are apparent at frequencies up to 16 MHz, the maximum frequency propagated by the 2-hop F mode. Between 16 and 26 MHz, the MUF, the bearings do not display large random fluctuations. Two interesting features occur in this frequency interval: (a) the sinusoidal like variations between 18 and 20 MHz, and (b) the sudden  $-1^\circ$  change at about 23 MHz.

A similar period to the previous example, but on a different day, is presented as Figure 4. In this case, a large systematic swing in bearing of about  $+7^\circ$  is evident between 10 and 22 MHz.

### 3.2 Diurnal effects

In order to examine the diurnal changes in both the ionograms and the associated bearing measurements, three types of graphical display have been developed. Examples of these are presented in Figures 5, 6 and 7 for the measurements of 8 January 1992.

The modal structure of the signals as a function of time of day and frequency is presented in Figure 5. In this diagram, each mode is represented by a different shade on a grey scale. The black regions indicate the frequency range over which both high and low angle signals are received via F-region propagation. This corresponds to the F-region 'nose' near to the MUF on the ionogram of Figure 4. Below this high/low angle region is the frequency range at which the signal is propagated by a single F-region reflection, and below this the region of multiple moded propagation, both by multiple F-region hops and by E-region reflections (the dash indicates the maximum frequency supported by the E-region). Of particular note are the presence of sporadic E (Es) reflections between about 03:00 UT and 09:00 UT which increase the

MUF to frequencies well in excess of those supported by the F-region.

Corresponding to the modal structure indicated in Figure 5, the signal strengths are presented in Figure 6, the darker shades representing the stronger signals. Of particular note is the low signal level of the sporadic E reflections, there being no apparent difference from the background signal level.

The bearing spreads (standard deviations) corresponding to the data presented in Figures 5 and 6 are given in Figure 7. The most striking feature of these measurements is the high correlation of the region of low bearing spread with the region of single moded propagation as indicated in Figure 5. Larger bearing standard deviations are observed when the signal is not single moded. Few good bearing measurements occur at frequencies less than about 6 MHz (the DF does not perform particularly well at these low frequencies).

#### 4. CONCLUDING REMARKS

A new technique has been developed to investigate the performance of an HF DF system as a function of frequency and mode structure and also to investigate the very large bearing errors which occur at high latitude sites. The system has been deployed at two mid-latitude sites and several days of data collected for paths to Ottawa and to Cheltenham. Examples of these preliminary measurements have been presented and several interesting features observed in the data identified. Further tests have been conducted at Alert, a very

high latitude site, during April 1992. However, at the time of writing, these additional measurements had not been analysed

#### 5. ACKNOWLEDGEMENTS

The authors would like to thank the Canadian Department of National Defence for their support of this project, GCHQ for access to their DF facilities and the DRA, Aerospace Division for the ROSE signal processing unit. PH would also like to acknowledge the SERC and DRA, Aerospace Division for a CASE studentship.

#### 6. REFERENCES

1. Jones, T.B. and Warrington, E.M. "Large bearing errors observed at a high latitude DF site", in "Radiolocation Techniques", AGARD Conference Proceedings, June 1992, paper 18.
2. Arthur, P.C., Dickson, A.H. and Cannon, P.S. "ROSE - a high resolution, amplitude coded, high frequency oblique ionosonde", in "Remote Sensing of the Propagation Environment", AGARD Conference Proceedings, September - October 1991, paper 10.
3. Warrington, E.M. and Jones, T.B. "Bearing determination and bearing quality indication from a goniometric HF DF system", in "Radiolocation Techniques", AGARD Conference Proceedings, June 1992, paper 20.

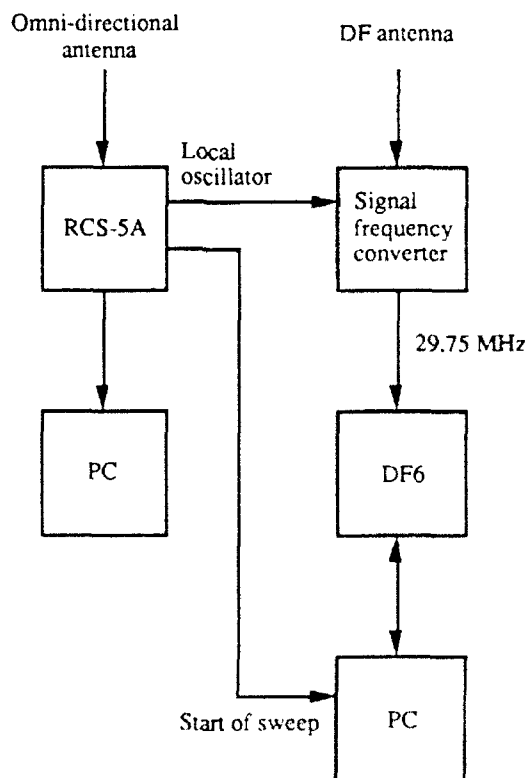


Figure 1. Block diagram of the sweeping DF experimental configuration.

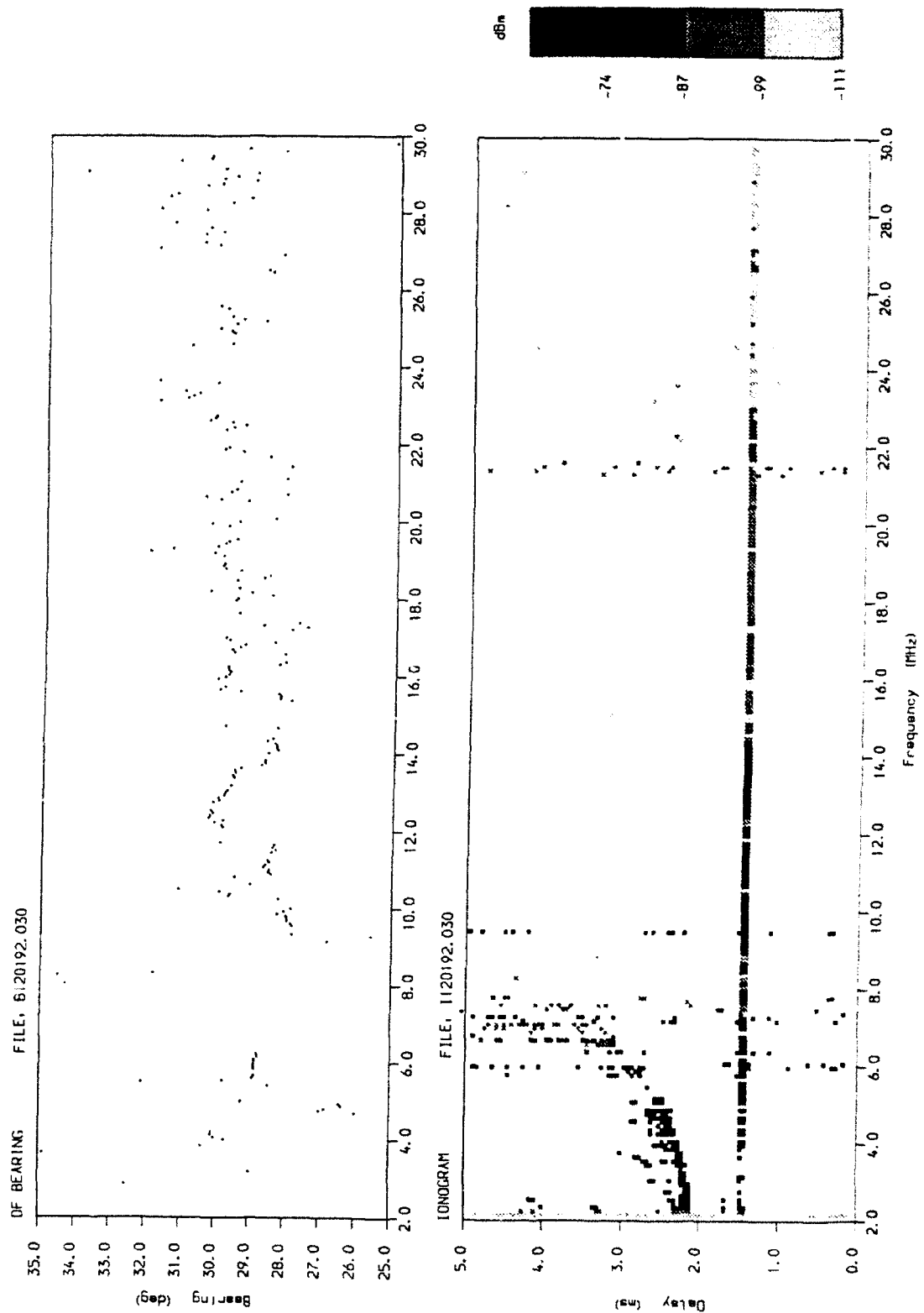


Figure 2. Oblique ionogram and associated bearing measurements for the Oslo - Cheltenham path. 02:30 UT, 12 January 1992.

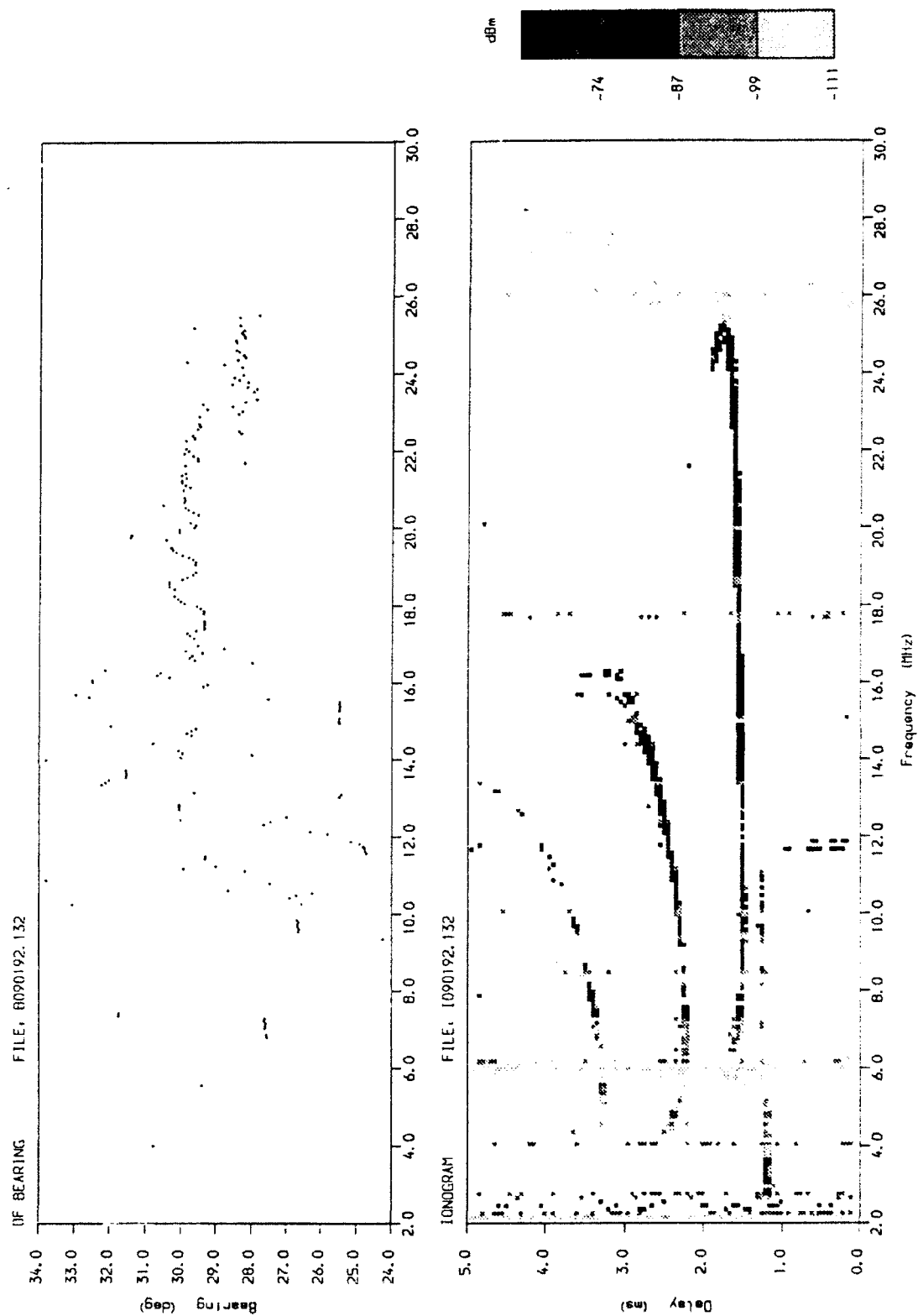


Figure 3. Oblique ionogram and associated bearing measurements for the Oslo - Cheltenham path, 11:00 UT, 9 January 1992.

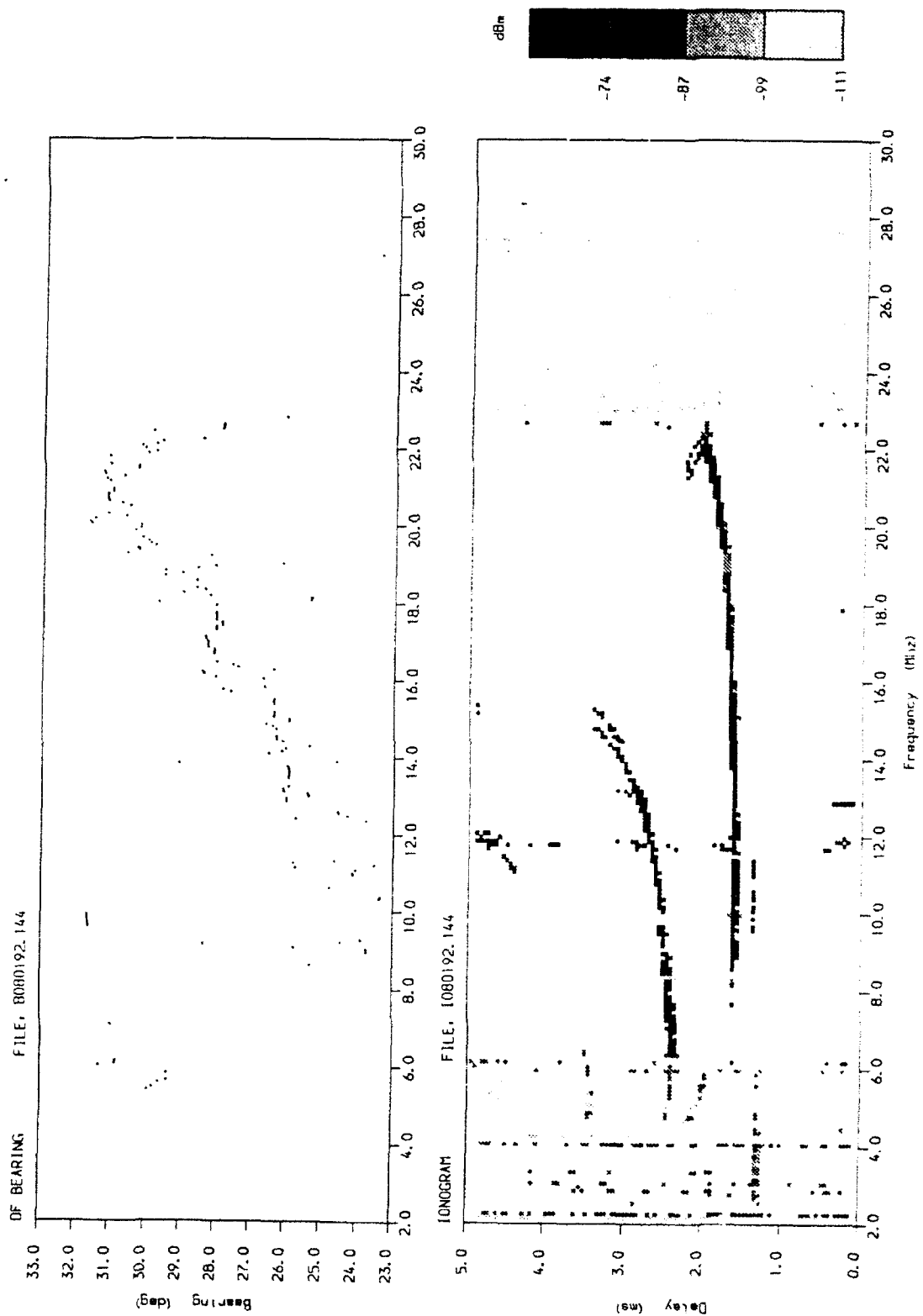


Figure 4. Oblique ionogram and associated bearing measurements for the Oslo - Cheltenham path. 12:00 UT, 8 January 1992.

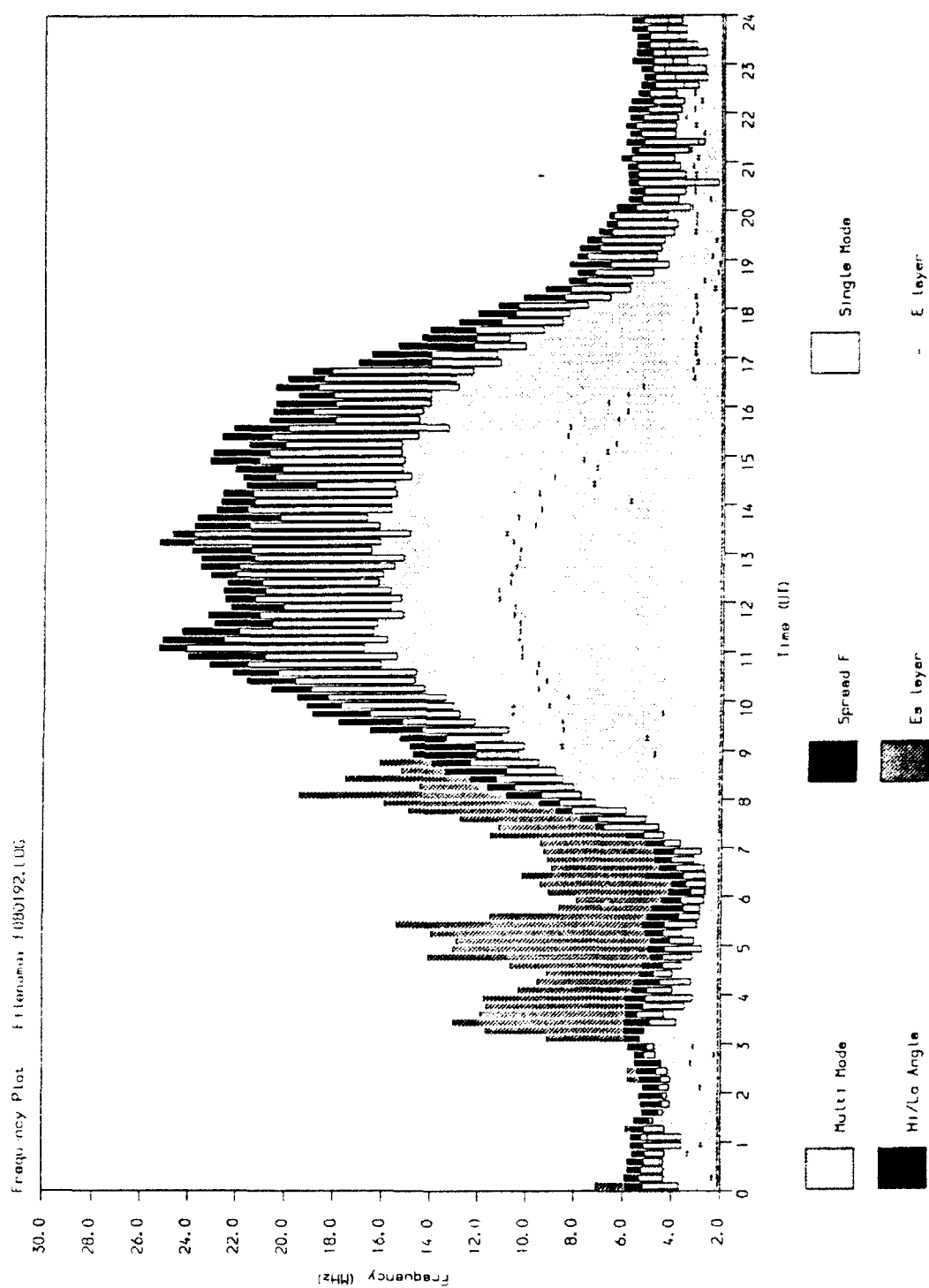


Figure 5. Mode structure determined from the oblique ionogram measurements for the Oslo - Cheltenham path. 8 January 1992.

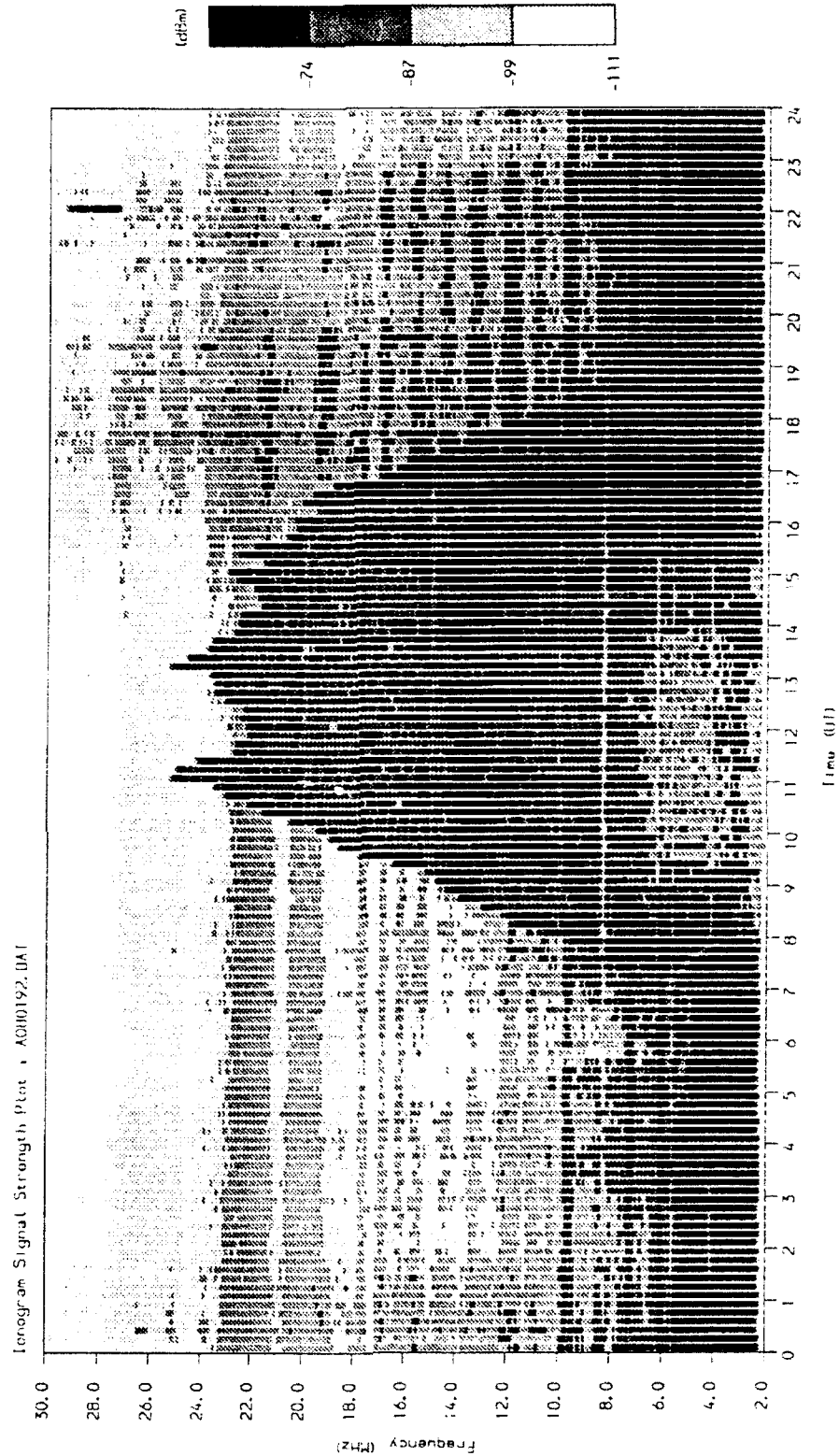


Figure 6. Signal strengths determined from the oblique ionogram measurements for the Oslo - Cheltenham path. 8 January 1992.



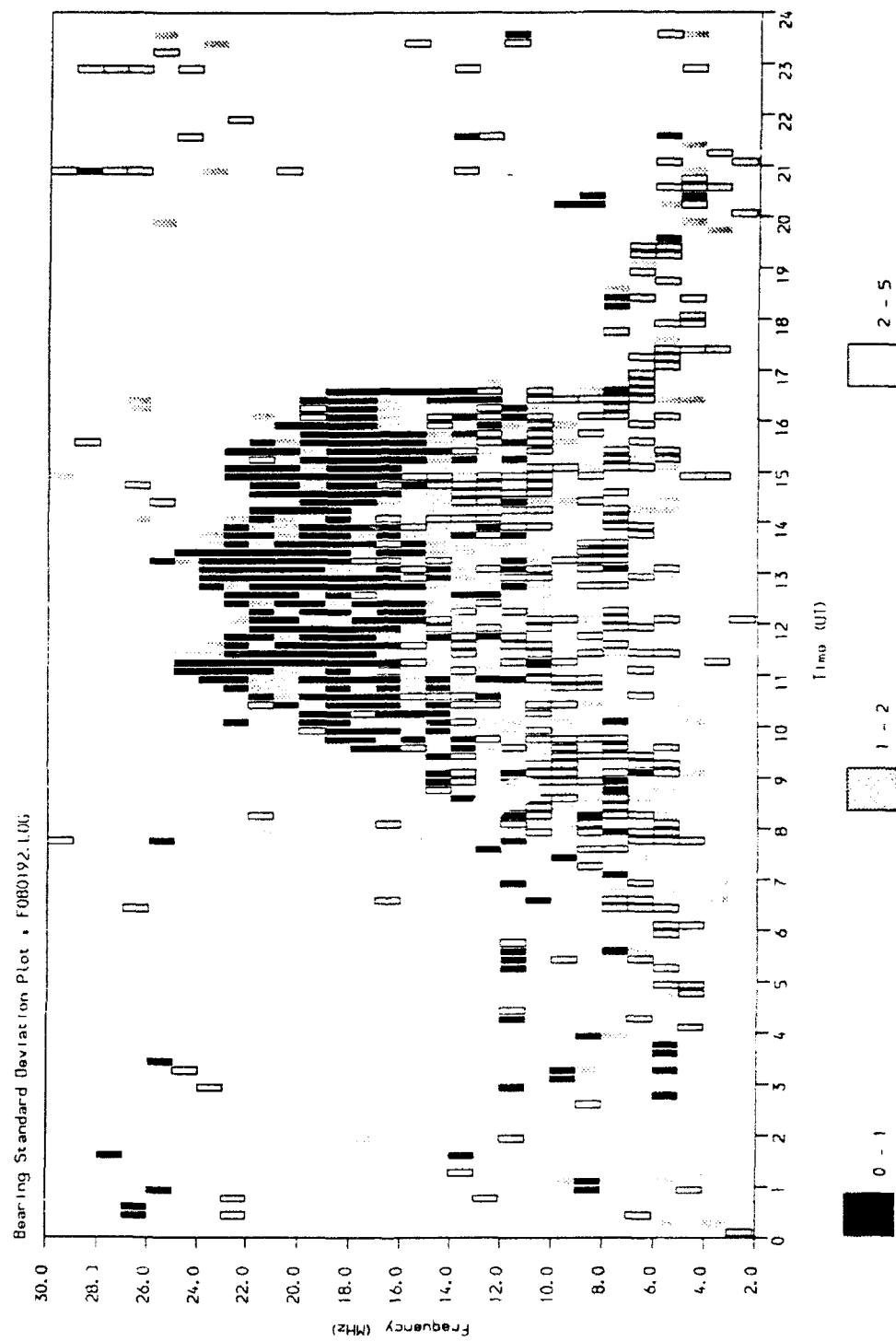


Figure 7. Bearing spread (standard deviation) measurements for the Oslo - Cheltenham path. 8 January 1992.

## DISCUSSION

### L. BERTEL

Your system determines only one azimuth angle - so it is necessary to be careful when you are in a situation of multihops, multimodes (o-x) or multipath: the azimuth angle calculated will vary with the doppler shift intermode as the elevation angle varies but with less amplitude (ref paper 9).

### AUTHOR'S REPLY

That is true when trying to determine the bearings associated with the individual modes. Our system does, however, give the bearings produced by the DF at each frequency with the prevalent modes. As such, it does give us an assessment of DF performance as a function of mode content and frequency.

### G. HAGN

1. Direction finding on the individual modes is an excellent idea, and you have shown some very promising results. The MUSIC algorithm, with the proper antenna (e.g., elements with a common phase center for the 3 orthogonal components of the arriving E and H fields), and a dechirping algorithm could be used to obtain simultaneously the azimuth and elevation angles of the modes which are resolved. This should work when several modes are present simultaneously - until the number of degrees of freedom required is exceeded. Getting the angles of arrival of the several modes simultaneously, would seem to be the next step in the evolution of what you have presented.

2. I believe your explanation of the reason for seeing the "weaker" high ray near the MOF, namely the difference in elevation plane patterns between your monopole and the Pusher with its wire ground plane. I have measured the elevation plane pattern of the 7.5 foot monopole with 16 ground radials 40 feet long and buried 7 inches, and the results agree well with computations made with the numerical electromagnetics code (NEC-3) - see IEEE Trans Broadcasting, No. 34, Vol. 2, pp. 221-229, June 1988. I also have made similar measurements on the Pusher (unpublished). Due to the extreme gradient of gain at low elevation angles for the monopole, it is important (if you can) to use a Pusher element (with a receiving multicoupler). Do you agree? Could you comment on the groundplane of your monopole?

### AUTHOR'S REPLY

1. We do hope to be able to develop our system along these lines in the future. However, at present, the only DF available to us is the goniometric DFG system.

2. We agree entirely. To be able to fully compare the DF and ionogram traces it is essential that a Pusher element be used with the sounder receiver since this will have the same vertical elevation angle sensitivity as the DF. Unfortunately, this was not available to us during the January test. A Pusher element was employed during our main experiment at Alert.

### W. SHERRILL

In an FMCW sounding, each mode appears in a different frequency bin of the spectrum-analyzed receiver passband. We have used conventional interferometry to measure the azimuth and elevation of each mode, assigning a DF process to each frequency bin (independently) which contains a mode.

Have you assigned your CDAA DF processes to individual mode frequencies?

### AUTHOR'S REPLY

No.

# Bearing determination and bearing quality indication from a goniometric HF DF system

E.M. Warrington<sup>1</sup>

T.B. Jones<sup>2</sup>

<sup>1</sup>Department of Engineering

<sup>2</sup>Department of Physics and Astronomy

University of Leicester

University Road

Leicester, LE1 7RH

UK

## SUMMARY

A computer based model of an automated wide aperture goniometric HF direction finder (type DF6) has been developed to investigate the system performance under various simulated signal environments. Several test runs of the model indicated that the original DF6 algorithm could be modified to yield a significant improvement in the measured bearing accuracy and in addition provide a meaningful indication of the quality of the measured bearing.

These modifications have been implemented and tested at two sites in Canada. A system located in Ottawa has been tasked against known transmissions from Halifax, Nova Scotia (950 km) and several hours of data collected. These measurements confirmed that the changes implemented as a result of the computer modelling produced improved bearing measurements. Further tests were conducted at Alert, a very high latitude site in the Canadian Arctic. At this site, the bearings often varied rapidly by several tens of degrees. These latter observations indicated that further changes to the DF algorithm should be made for systems deployed at this high latitude site.

## 1. INTRODUCTION

The performance of HF direction finding systems is difficult to evaluate solely from measurements of real signals propagated via the ionosphere due to several factors. Principally, the true bearing may not be known since the signal may arrive at the DF over off-great circle paths due to the presence of ionospheric tilts and, furthermore, the bearing may not remain constant due to changes in the ionospheric tilts.

To overcome these difficulties, a computer model of a wide aperture type DF6 direction finder has been developed. By means of the model, the performance of the DF system has been assessed under well defined conditions and the affect of changes to the bearing calculation algorithm determined. In addition to calculating the bearing, parameters of the received DF waveform can provide an indication of the bearing quality (likely accuracy).

Note: In this paper the term 'bearing' refers to a measurement of the azimuthal angle of arrival of the target signal. Ionospheric effects may be present which result in the measured bearing not being along the great circle path between the transmitter and receiver. A bearing which is referred to as accurate or of high reliability is intended to indicate that the measurement of the azimuthal angle of arrival is accurate or of high reliability. No consideration of ionospheric effects leading to off great circle angles of arrival and hence to line of bearing errors has been made.

## 2. THE DF6 SYSTEM

The DF system employed in this investigation is a wide aperture goniometric system with a dual band antenna array, a single receiver and a computer based data collection and processing system. A brief description of the principal components and the method of operation is presented below.

The antenna array (a PUSHER) consists of two concentric 24 element circular arrays with diameters of 75 m and 25 m. The larger of the two arrays (the lowband array) operates at frequencies of between 3 MHz and (usually) 10 MHz, and the smaller of the arrays (the highband array) at frequencies between (usually) 10 MHz and 30 MHz. At Ottawa and Alert, both of the arrays consist of single element vertical monopoles.

RF signals from the sum beam output of the PUSHER antenna are demodulated in the system receiver and combined with goniometer positional information for initial processing. An average beam pattern, known as a composite scan, is then formed from successive rotations of the goniometer. Line bearings are then computed by applying software algorithms to the digitised composite goniometer sum beam patterns.

In the original algorithm, 16 composite scans are input and the following procedure adopted to calculate the bearing(s) of the incident signal(s):

1. For each composite scan, a snap bearing is determined as the point of best symmetry of the DF waveform. This is found by folding the DF waveform at the peak value and

moving the fold point so as to minimise the area between the original and folded curves (see Figure 1). If a minimum does not exist within  $\pm 12^\circ$  of the peak, the snap is flagged as invalid.

2. In order to separate snap bearings obtained from signals arriving from different azimuth angles, all valid snaps are then allocated to a series of  $10^\circ$  wide overlapping bins (see Figure 2). Overlapping bins are employed in this procedure to prevent problems arising when points are distributed around a bin boundary causing them to fall into adjacent bins. For the situations where several signals are present, the system relies on the presence of modulation and fading to ensure that significant numbers of snaps are determined for each of the constituent signals.

3. For the bin containing the most snaps, if four or more, a bearing is calculated as the mean of the points within that bin. These snaps are then removed from that bin, and also from the overlapping bins into which they were allocated.

4. A quality factor is then calculated as  $20\sigma^2/(n-1)$ , where  $\sigma$  is the standard deviation and  $n$  the number of points within the bin. If the calculated value exceeds 99, then the quality factor is set to 99. High values are intended to indicate bearings of poor reliability, and low values bearings of high reliability. In this context, a high reliability bearing refers to an accurate measurement of the azimuthal angle of arrival of the signal. Ionospheric effects may mean, however, that this is not the great circle direction.

5. The procedures outlined in (3) and (4) are then repeated until no bin contains four or more points.

Experimental measurements indicate that the quality factor calculated by this method is not meaningful. An evaluation of this parameter for the 8.697 MHz transmission from Halifax, Nova Scotia received at Ottawa is presented as Figure 3. In this diagram, each bearing returned by the DF is shown as a point on a scatter plot where the x-axis is the measured bearing and the y-axis the quality factor. Most of the points for this example have quality factors less than 10, and measured bearings in the range  $85^\circ - 93^\circ$ . No indication can be obtained from the quality factor as to which are the more accurate measurements.

### 3. COMPUTER MODELLING

By means of the computer model, the performance of the DF system was evaluated for a range of realistic signal environments. The model allows for several incident wavefronts at specified directions of arrival (in elevation as well as azimuth) with various frequency differences (analogous to Doppler shifts for different modes of the same signal or differences in operating frequency for interfering signals). Each of the components can be allowed to fade, the variations in amplitude and phase being generated from the model given by Gething [1]. A specified level of random noise can also be added.

The model produces samples of the receiver output (either the sum or difference beam pattern) at  $2^\circ$  intervals in the goniometer rotation. These samples then form the input to a model of the DF system. In the case of the original DF6 algorithm, averages of the model output for four goniometer

rotations are produced. These averaged beam patterns (composite scans) are then input to a third procedure in which the algorithms employed by the DF6 are applied to the data.

As a result of the modelling exercise, the following enhancements were made to the algorithm which resulted in a better performance (of the modelled system) when compared to the original algorithm.

#### 3.1 Windowing

To calculate a bearing from the goniometer output pattern, the point of best symmetry of the pattern is obtained by the folding process described in section 2. The presence of additional peaks or distortions in the goniometer output pattern due to the presence of several modes of propagation or interfering signals can severely affect the azimuthal angle for which the best point of symmetry is found. Computer simulation of this technique indicates that better (more accurate) results can be achieved by reducing the azimuthal range of the DF waveform included in the folding process (see Figure 4). The optimum azimuthal range, referred to as the folding window, is approximately the width of the main peak in the goniometer output pattern for the case when a single signal is present. Note that in selecting the window width, a slightly narrow window (ie. narrower than the main peak) has a much worse performance than one slightly wide. For this reason, the window width was selected to be 1.1 times the width of the main lobe for an elevation angle of  $20^\circ$ .

#### 3.2 Asymmetry

The quality of the folding process can be ascertained from the area between the two curves defined by the fold (see Figure 4) within the azimuthal window described above. For perfect symmetry this area would be zero. Since the difference in area is influenced by the signal amplitude, normalisation is achieved by dividing the area between the curves by the total area under the curve within the folding window.

Two typical modelled situations for a 10 MHz signal, bearing  $180^\circ$  with signal to noise ratios of -10 dB and +10 dB and the presence of several interferers is presented as Figure 5. In this diagram, snap bearings are shown as points on a scatter plot of asymmetry vs bearing. From this diagram it is apparent that snaps with low asymmetry values have a lower azimuthal spread than those points at higher asymmetry values. The width of the spreading is approximately proportional to the width of the main lobe of the DF waveform.

### 4. OBSERVATIONS AT ALERT

In addition to the computer modelling exercise outlined above, bearing measurements were made at various sites, including Alert, located in the northern Arctic. At this northerly site, rapid fluctuations in bearing of several tens of degrees are observed, indicating that the  $10^\circ$  bin width employed in the snap editing process to form a composite bearing is insufficient at this site.

### 5. NEW ALGORITHM

The computer modelling and the observations from Alert suggest two improvements to the original DF6 algorithm which would result in an improved performance. These are outlined below.

### 5.1 Determination of snap bearings

Snap bearings are calculated for each composite scan by the folding process described above with the addition of the azimuthal window suggested by the computer based modelling. The width of this window is selected to be 1.1 times the width of the main peak in the DF waveform as calculated for an elevation angle of  $20^\circ$ . This increased window width allows for the slight broadening of the peak at higher elevation angles. The asymmetry value discussed in section 3(c) is also calculated for each snap bearing.

### 5.2 Calculation of bearings and quality factor

The procedure adopted for calculating the composite bearings to be returned by the system is similar to that employed in the original algorithm, with the following differences:

1. The asymmetry values associated with each snap bearing are used to allocate weighting factors to each of the snaps. These weights are approximately equal to the reciprocal of the square of the expected error of snaps with the specified asymmetry value as derived from the computer modelling. Since the spread in the snap bearing distribution is dependent upon the width of the main lobe in the DF waveform, the weights assigned are dependent upon the frequency of the signal and account is taken of this variation.
2. The snaps are allocated to bins as in the original algorithm and the same procedure followed except that the returned bearings are calculated as the weighted mean of the snaps within the bin.
3. The quality factor is calculated as  $1000/\sqrt{S}$ , where  $S$  is the sum of the snap weights. Since the weights are related to the expected errors of snaps with that weighting factor, the quality factor should be a measure of the likely error associated with the returned bearing. This quality indicator is normally allowed to take a maximum value of 99.
4. Optionally, the quality factor may be returned on a logarithmic scale as  $10\log(1000/\sqrt{S})$ .
5. The width of the bins employed in the snap editing process can be changed from the default value of  $10^\circ$  to allow for the large, rapid bearing fluctuations observed at Alert.

Illustrated in Figure 6 are scatter plots of composite bearing and logarithmic quality factor for the data of Figure 5. For the case of a low signal to noise ratio ( $-10$  dB), moderately spread bearings are calculated with quality factors between 20 and 30, indicative of poor quality measurements. For the higher signal to noise ratio case ( $+10$  dB), some of the bearings are much less spread and these have been allocated lower quality factors indicative of good quality measurements.

## 6. EXPERIMENTAL MEASUREMENTS

In order to test the new algorithm, several hours of data were collected at two sites in Canada. The first, Ottawa, is a mid-latitude site and the second, Alert, is located at a very high latitude ( $82.5^\circ$  N) in the North West Territories.

### 6.1 Ottawa

At this site, measurements were made of the CFH CW Morse code transmissions from Halifax, Nova Scotia (bearing  $90.2^\circ$ ,

range 960 km) on frequencies of 6.430 MHz and 8.697 MHz. Examples of these measurements are now presented.

#### 6.1.1 Windowing

The affect of applying a window in the folding process for the determination of snap bearings is illustrated in Figure 7(a) and (b). In this diagram, the time variation of snap bearings is presented over two 30 second periods for window widths of  $360^\circ$  (no window) and  $60^\circ$ . The variation in the snap bearings with time is much less when the folding window is applied than for the case when all of the goniometer output is included. Note, however, that the true bearing may not be constant due to ionospheric effects.

Figure 8 is a scatter plot of snap bearing vs asymmetry with a  $60^\circ$  folding window applied for the CFH transmission for the 30 second period commencing at 22:29 UT. The scatter of the data points agrees moderately well with that expected from the computer modelling. At other times the agreement is not always good. In general, however, it would appear that improved bearings are obtained by applying the azimuthal window in the folding process, and by giving more weight to snaps with low asymmetry values.

#### 6.1.2 Composite bearings

Figure 9 represents a scatter plot of bearings and associated quality factors obtained with the new algorithm for the 8.697 MHz CFH transmission. The bearings are spread over a range of approximately  $8^\circ$  with quality factors in excess of 30, indicative of poor quality bearings. The peak of the bearing distribution occurs at  $87.5^\circ$ , approximately  $2.7^\circ$  less than the great circle bearing to this transmitter. These data correspond to those presented in Figure 3 for the original algorithm. The distribution of bearings obtained from the original algorithm is spread over some  $10^\circ$ , with an ill defined peak at around  $86.5^\circ$ .

### 6.2 Alert

Tests of the new algorithm have also been undertaken at Alert, located high in the Canadian Arctic. At this site, rapid fluctuations in bearing of several tens of degrees are observed due to propagation effects. In this case more accurate composite bearings are determined by not weighting the snap bearing measurements in the editing process.

## 7. CONCLUDING REMARKS

Two modifications to the DF6 algorithm have been suggested as the result of computer simulations. The first of these, the introduction of an azimuthal window into the folding process to determine snap bearings, resulted in less random fluctuations of the measured bearings and consequently lead to improvements in the composite bearings. The second modification, the use of the beam asymmetry as an indication of snap bearing quality, resulted in the generation of a meaningful indication of the overall bearing quality. It should however be noted that this latter parameter requires the antenna / goniometer system to generate a perfect beam pattern when illuminated by a planar wavefield otherwise the distortions introduced by the system adversely affect the quality factor.

## 8. ACKNOWLEDGEMENTS

The authors would like to thank GCHQ and the Canadian Department of National Defence for their support of this work.

## 9. REFERENCES

1. P.J.D. Gething. Radio direction finding and the resolution of multicomponent wavefields. Peter Peregrinus Ltd. on behalf of the IEE.

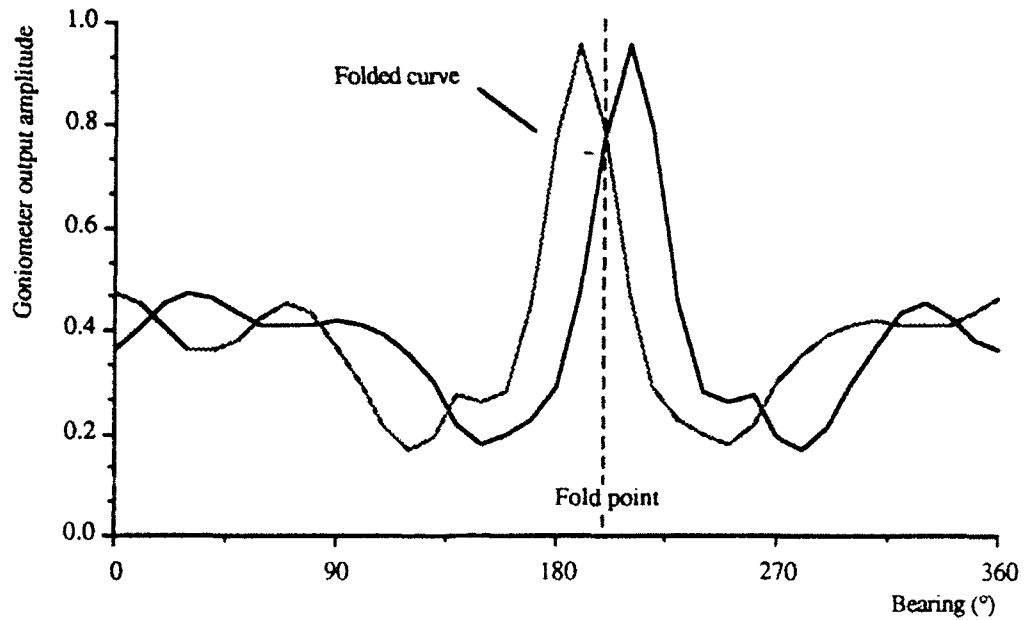


Figure 1. Illustration of the folding process employed to determine the snap bearing as the point of best symmetry. The fold point is moved such that the area between the original and folded curve is minimised.

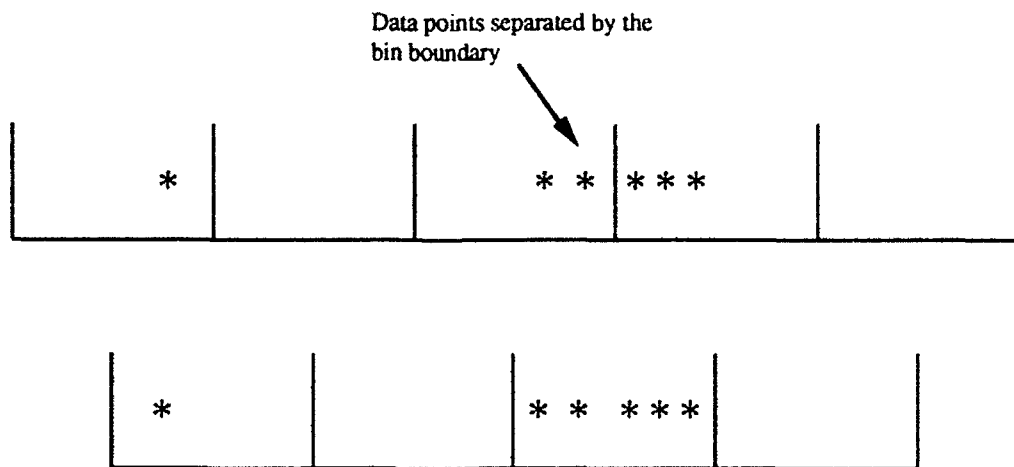


Figure 2. Allocation of the snap bearings to overlapping bins. The use of overlapping bins prevents problems arising when the true bearing is near to a bin boundary causing snaps to fall into adjacent bins.

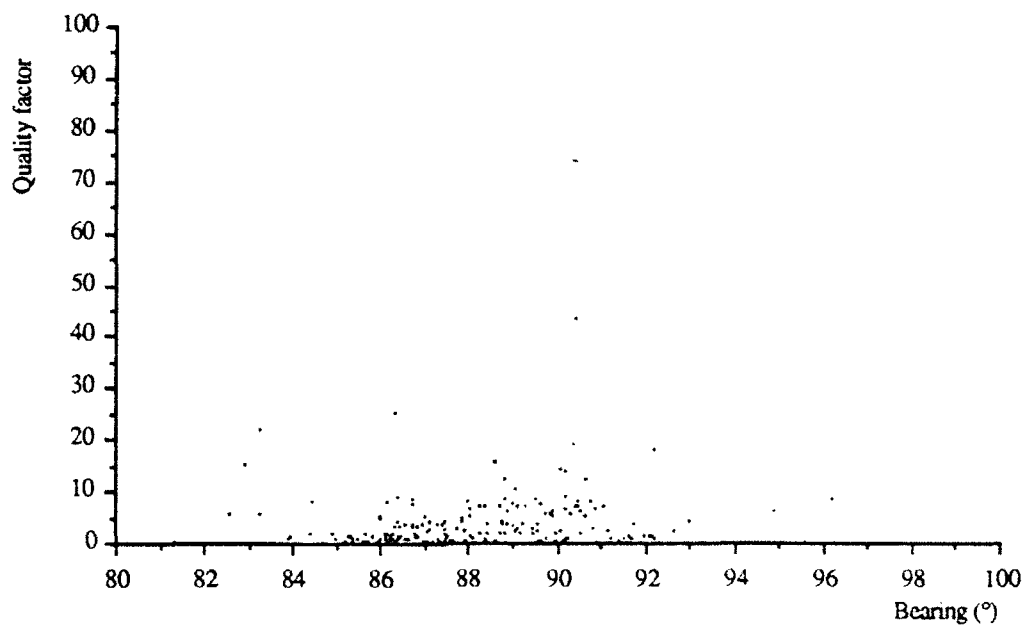


Figure 3. Scatter plot of bearings and associated quality factors obtained using the original DF6 algorithm for the 8.697 MHz CFH transmission from Halifax, Nova Scotia (true bearing 90.2°)

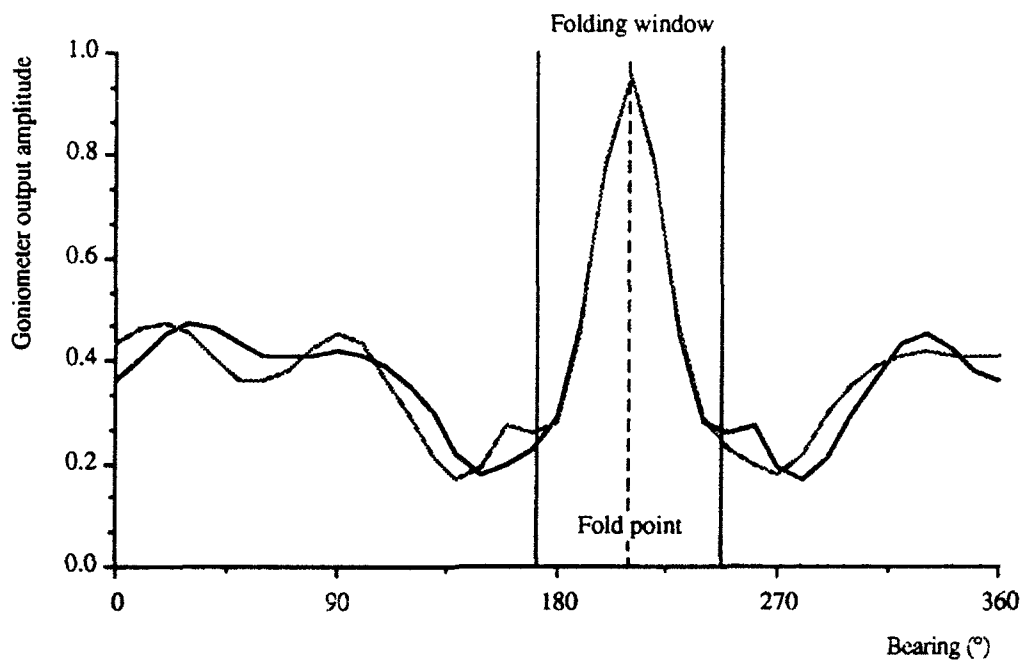


Figure 4. Illustration of the azimuthal window introduced into the folding process. In this case, only those parts of the original and folded curves within the azimuthal window are taken into account when determining the snap bearing.



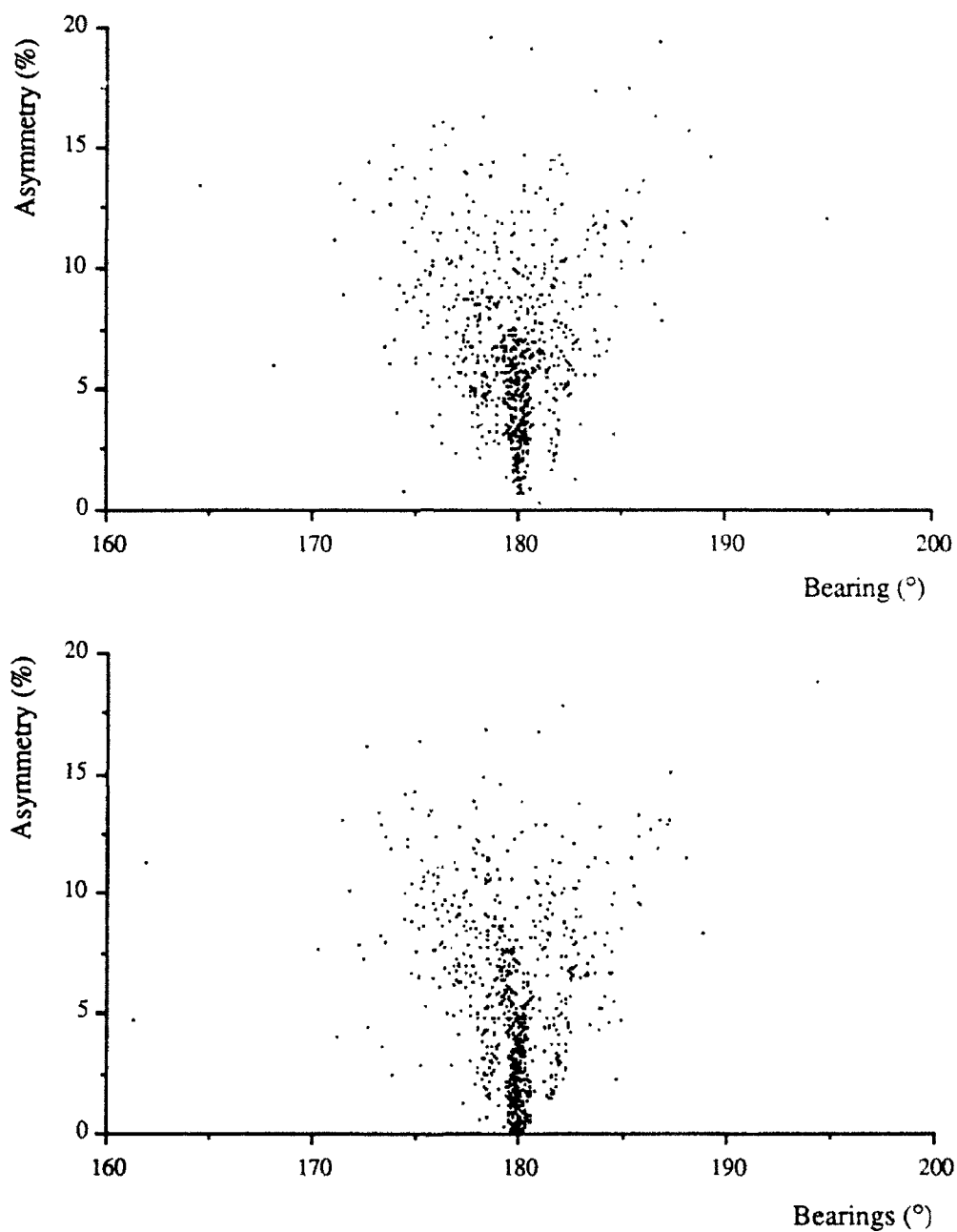


Figure 5. Modelled situation for a 10 MHz signal using the high band antenna with a 48° folding window applied. The signal has a bearing of 180° together with interfering signals of lower amplitude well separated in azimuth. The upper frame is for a signal to noise ratio of -10 dB and the lower frame for a signal to noise ratio of +10 dB

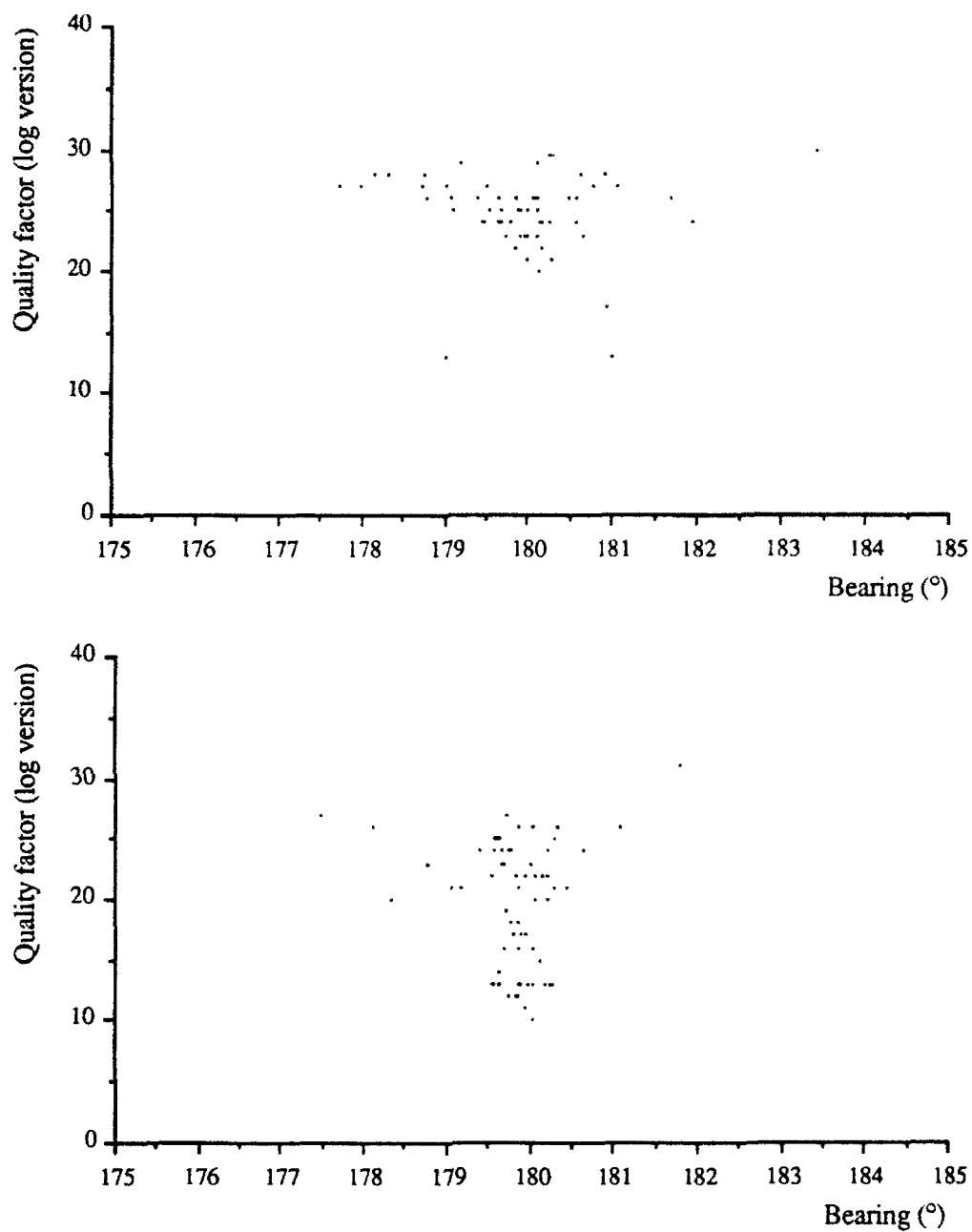
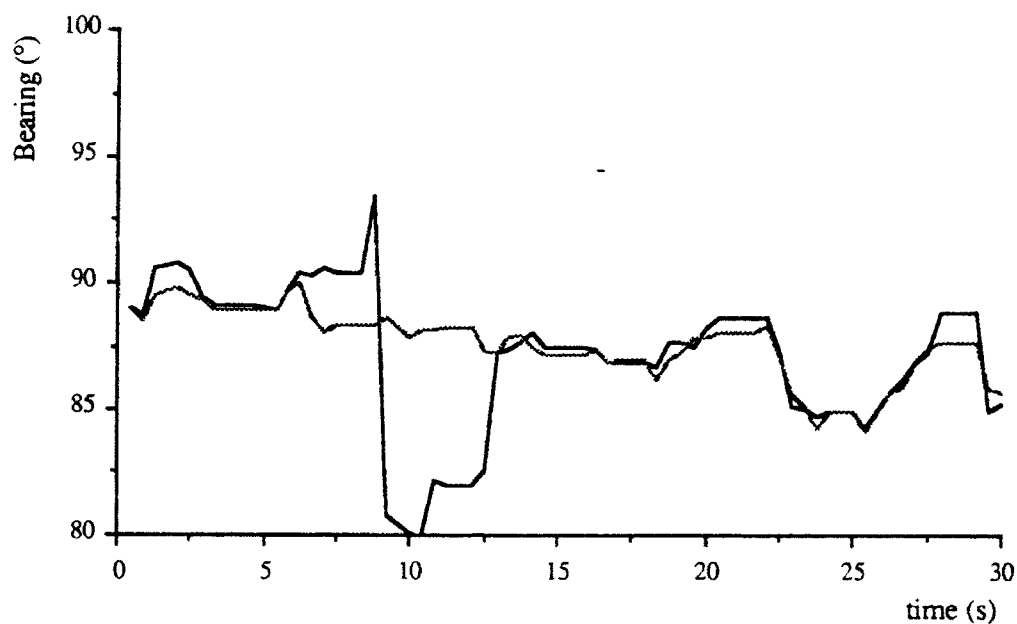
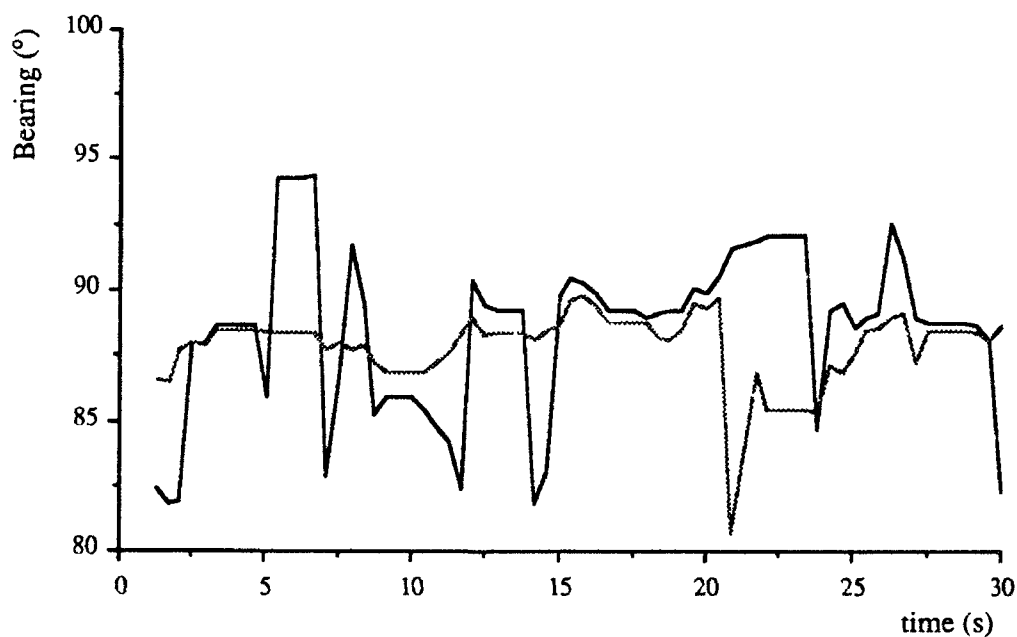


Figure 6. Scatter plots of the quality factor (logarithmic version) vs composite bearing for the data shown in Figure 5. 8 seconds integration time employed. The upper frame is for a signal to noise ratio of -10 dB and the lower frame for a signal to noise ratio of +10 dB



(a) 21:27 UT, 21 March 1989



(b) 21:33 UT, 21 March 1989

Figure 7. Plots of snap bearing vs time for two 30 second periods for the 8.697 MHz CFH transmission. No windowing has been applied in the snap bearing determinations indicated by the dashed lines, whereas a 60° window has been applied for the bearings indicated by the solid lines.

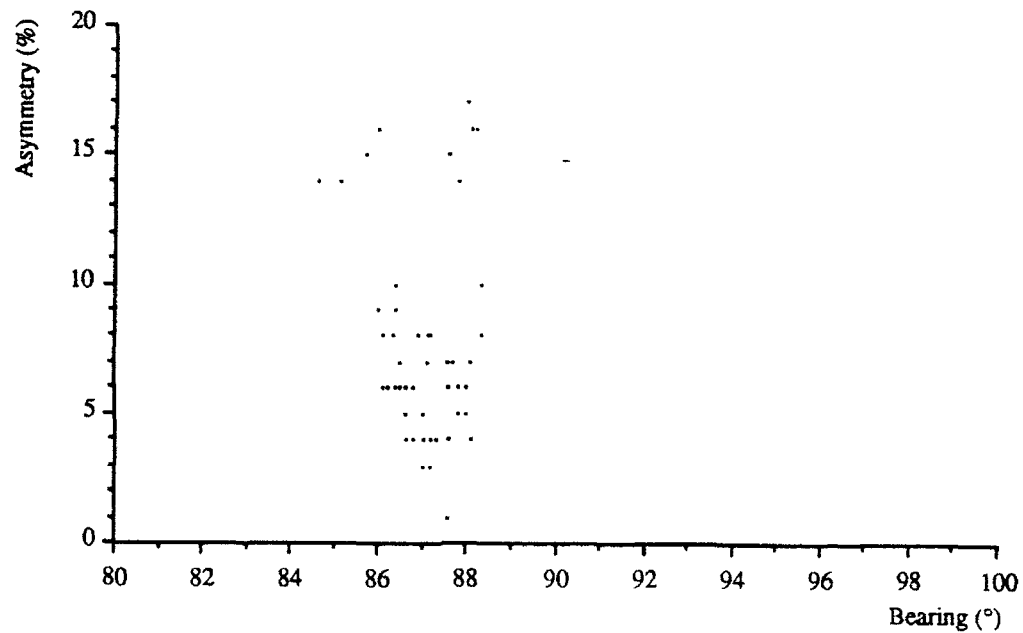


Figure 8. Plot of snap asymmetry vs bearing for the 8.697 MHz CFH transmission. Data are for the 30 second period starting at 22:29 UT, 21 March 1989

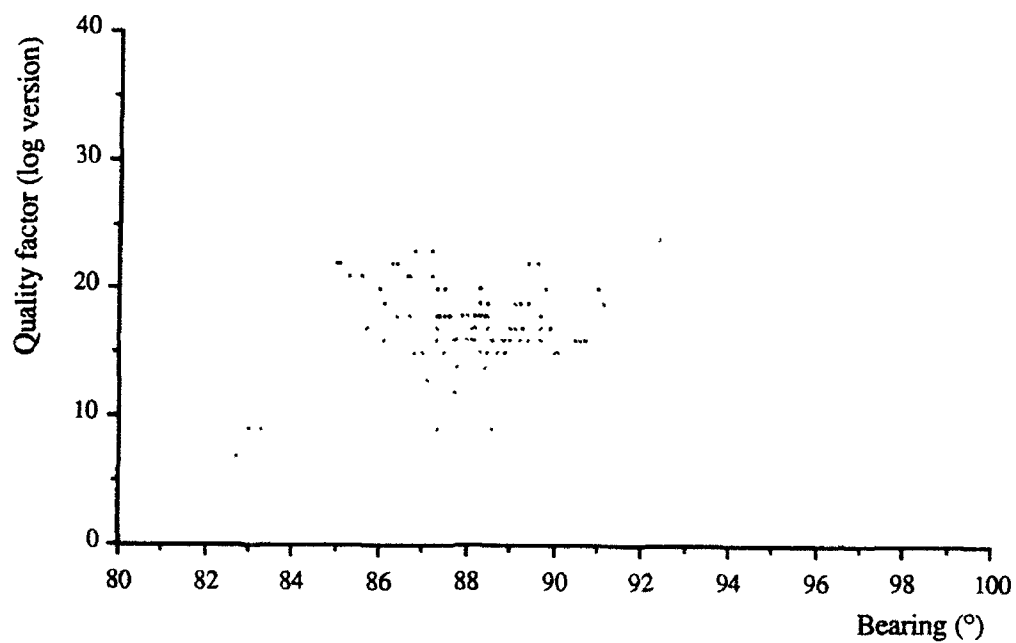


Figure 9. Scatter plot of bearings and associated quality factors obtained using the new algorithm for the 8.697 MHz CFH transmission. These data correspond to the data presented in Figure 3 for the original DF6 algorithm.

## DISCUSSION

### J. BENDER

How did you define signal-to-noise ratio and how did you measure it? In similar trials in Germany, we had problems in measuring S/N because of its variability.

### AUTHOR'S REPLY

Signal to noise ratio was not measured during the system trials, but was one of the parameters used by our computer model. A random vector was added to the signal vectors in the model and its average power was determined by the specified SNR.

## ANGLE OF ARRIVAL CHARACTERISTICS OF IONOSPHERIC SKYWAVE SIGNALS

D. Mark Haines  
Bodo W. Reinisch  
University of Massachusetts Lowell  
Center for Atmospheric Research  
450 Aiken Street  
Lowell, MA 01854  
U.S.A.

### SUMMARY

Skywave signals in the HF band experience a plethora of effects which distort them in the time, space and spectral domains. The ability of a particular measurement to resolve discrete components which faithfully represent a single signal source (i.e. useful for angle-of-arrival estimation) is a somewhat random occurrence given any one fixed integration time. However, the fact that the signal components can be resolved in the spectral domain by our Doppler technique much of the time seems to indicate that there are indeed a finite number of discrete components in any skywave signal and that the Doppler shift for each component varies independently such that they may be resolvable in the spectral domain.

### INTRODUCTION

Observation of the ionosphere with ground based sensors relies on the accurate measurement of observable features of an electromagnetic wave transmitted through the ionospheric medium. These observables are:

1. frequency
2. height or range (i.e. time delay)
3. amplitude
4. phase
5. Doppler shift or spread (frequency dispersion)
6. wave polarization (left or right hand circular)
7. angle of arrival

Given the ability to measure all of these parameters one can describe the vertical electron density profile, the existence direction and velocity of wave structures (like the traveling ionospheric disturbances associated with acoustic gravity waves) /1/ and plasma irregularities carried by ionospheric drifts and polar cap convections /2,3,4/. The accurate measurement of all of the above parameters, except frequency, depends heavily on the signal to noise ratio of the received signal. Therefore vertical incidence ionospheric sounders capable of acquiring high quality scientific data have historically utilized powerful pulse transmitters in the 2 to 30kWatt range.

The necessity for extremely good signal to noise ratio is demanded by the sensitivity of the measurements to variations in the signal level. For instance, to measure phase to 1 degree accuracy requires a signal to noise ratio better than 40dB (assuming a Gaussian noise distribution which is actually a best case), and measurement of amplitude to 10% accuracy requires over 20dB signal to noise ratio. Of course, we also would like these measurements to be immune to degradation from noise and interference and to maintain their high quality over a large frequency band which requires that at the lower end of the HF band we have to overcome absorption, noise and interference and still provide at least a 20 to 40 dB signal to noise ratio. Therefore the goal of making an inexpensive low-power portable sounder would at first seem unattainable.

### DESCRIPTION OF THE DIGISONDE PORTABLE SOUNDER

#### Physical

The Digisonde Portable Sounder (DPS) is for the most part a miniaturized implementation of the well proven measurement techniques used by the Digisonde 256 sounder /5,6/ produced for the past 10 years by the University of Lowell Center for Atmospheric Research (ULCAR) (now through change of name only, the University of Massachusetts Lowell) Center for Atmospheric Research. The addition of digital pulse compression makes the use of low power feasible and the high level language control software and standard PC-DOS (i.e. IBM/PC) data file formats provide a new level of flexibility.

Figure 1 shows the DPS system. The upper chassis contains the system computer, signal generation, reception and processing hardware and over 200 Mbytes of data storage capacity, the lower unit contains a 500 Watt solid state amplifier. The two 19" rack mountable chassis are installed in a lightweight fiberglass transport case which includes a shock mounted 19" rack. The components in the fiberglass case are, by themselves (with the addition of a 24 volt battery and antennas) capable of making and storing preprogrammed automatic measurements. The peripherals (keyboard, monitor, printers and modem) are only required for human interface.

#### Signal Processing

The DPS is able to be miniaturized by lengthening the transmitted pulse of a small low voltage solid state transmitter we can transmit an amount of energy equal so that we transmit by a high power pulse transmitter without having to provide components to handle the very high voltages required for 10's of kilowatt power levels. The time resolution of the short pulse is provided by intrapulse phase modulation using one of three programmable phase codes. The use of a Complementary code /7,8/ pulse compression technique is described in our previous paper /13/, which shows that the expected measurement quality is the same as a conventional sounder of 1MWatt peak pulse power.

#### Multiplexing

The Digisonde Portable Sounder allows multiplexed Doppler integration of up to 64 separate combinations of frequencies, antennas (the system includes 4 phase matched antennas switched into 1 receiver, while the newer DPS-4 contains 4 receivers to allow "true" simultaneous reception on the 4 antennas), and polarizations (i.e. Ordinary or Extraordinary circular polarization). This multiplexing is implemented by switching hardware parameters between pulses, and is useful for making high Doppler resolution ionograms, making precision group height measurements and accurately determining O or X polarization as described in /13/. The angle of arrival measurement of the DPS-1 (the system

which produced all the data presented here) is made possible because the multiplexing allows the signals on the four antennas to be integrated simultaneously, therefore providing phase differences in accordance with the received signal's angle of arrival. This essentially simultaneous (actually interlaced or multiplexed) reception is accomplished by synchronizing the computer's control, data acquisition and signal processing software functions to the transmitted waveform through a hardware interrupt. Two systems working in a bistatic mode can be synchronized by locking that hardware interrupt to the 1 PPS (pulse/sec) signal from a GPS satellite receiver mounted in the DPS system. At the end of a multiplexed Doppler integration the system will contain an entire Doppler spectrum for each receiving antenna, each frequency and each polarization measured, at each resolvable range. For instance a 40 second coherent integration could contain up to 8192 128-pt Doppler spectra any or all of which may be stored on a 150MByte tape cartridge. If multiple sources exist at the same range (height for vertical measurements) they are usually resolved in the Doppler spectrum computed for that height. Since the four antenna signals were integrated simultaneously, the phase of each of these complex Doppler lines will now yield the true angle of arrival by simple interferometry (triangulation) techniques. The fourth antenna provides a consistency check to detect any Doppler lines which still contain multiple signal sources.

#### Automatic Sequencing

Table 1 shows the DPS measurement menu. In our previous paper /13/, the function of each item in the menu is explained, but a summary of its use follows. Parameters 1, 2 and 3 specify the frequency range step size over which the range and step size measurement is made. The next 4 parameters set up the multiplexing of frequencies, transmitter phase codes, antennas and polarization. The next 3 parameters determine the pulse repetition rate, number of pulses to be integrated and sample spacing (in km). The last 5 parameters determine the format in which data will be output to hard disk, 150MByte tape cartridges or printers. Since there are up to 8192 Doppler spectra to store and typically less than 1% of this is useful data, much of the science in making various measurements is affected by this choice of data format. The trade-off depends on the physical phenomena being studied and on the allowable frequency of changing the data recording media, but recording all data is one option.

Table 1. Program Editing Screen

(L) Lower Freq (kHz) 2000	(H) Height Res (km) 5
(C) Coarse Step 100	(M) # of Hgts 128
(U) Upper Freq (kHz) 12000	(D) Delay (1=50μS) 0
(F) Fine Freq Step (kHz) 5	(G) Gain (0 to 15) 9
(S) # Small Steps 2	(I) Freq Search (0,1) 1
(X) Xmitr Waveform 1	(O) # Output Hts 8
(A) Antennas (0=Beam) 0	(P1) Disk (0,A,M,D,F,P) A
(N) FFT Size (Power of 2) 5	(P2) Printer (BW, Color) 0
(R) Rate (PPS) 200	(B) Bottom Ht 210
(E) First Height (km) 90	(T) Top Ht 500

#### EXPERIMENTAL DATA SHOWING AOA CHARACTERISTICS OF SKYWAVE SIGNALS

There are many propagation phenomena involved in ionospheric skywave propagation which result in the appearance of several sources at the receiver from a single transmitter location. For instance, for oblique propagation characterized by a parabolic electron density profile there is the splitting into two rays, a high angle and a low angle ray /10/. Then there are multihop paths, O and X polarization modes, and ionospheric roughness (irregularity structures) which provide multiple propagation paths between transmitter and receiver.

Application of interferometry principles to signals received on three or more physically separated antennas can accurately determine the signal's angle of arrival. IF AND ONLY IF only one source (i.e. one direction of arrival) is involved in creating the received signal whose phase is measured. Unfortunately the mechanisms which create multiple source locations in the skywave signal result in multiple superposed signals at the receiver. The phase of such a signal is the geometric vector sum of all of the separate source components and is not characteristic of any one direction of arrival. If a means of separating these signal components (e.g. resolution in range or Doppler) can be implemented, such that only one source exists in a given range/Doppler cell then the phase of that component on each of several (or at least three) antennas can be used to determine its angle of arrival.

#### Angle of Arrival Spreading for Vertical/Near Vertical Incidence

ULCAR's Digisonde sounders (the DPS and previous Digisonde versions) use a processing program we call ODDA (Online Digisonde Drift Analysis) to compute and display these angles of arrival along with the 3-dimensional position and velocity (speed and direction of motion) of plasma features creating the echoes. The DPS can simultaneously measure up to 16 frequencies using each of the four receive antennas while recording echoes from 128 heights (with 5km height sample spacing this covers 640km of range).

Figure 2 shows the screen display for a four frequency measurement. The first four vertical height profiles show echoes for the four different frequencies coming from 330, 335, 340, and 345km (referenced to the axis in km on the left side). The next four vertical strips show the contents of the 128-point Doppler spectrum at the height of maximum echo amplitude (referenced to the axis in Hz on the right side), which in this case is the only echo height (270, 225, 280 and 285km). This 20 second integration can be repeated an indefinite number of times to allow observations as a function of time; 8 repetitions are shown in this example. The entire Doppler spectrum at each of the echo heights is stored as input for the ODDA program.

The ODDA program /12/, after screening the data with several quality and consistency checks, uses the 4 recorded phases (one for each antenna) at each of the Doppler lines to compute a 2-Dimensional angle-of-arrival map (a "skymap") for each height recorded. Doppler integration times of 10 to 40 seconds are usually sufficient to ensure that each Doppler line contains only a single source component, but a consistency check using the fourth antenna can detect and exclude points resulting from a non-planar wavefront. Echo locations can be accumulated on the plot for a selectable period of time, but too long a time may obscure some of the short term phenomena being observed as happens in Figure 3 taken in March 1992 by the DPS system in Jicamarca, Peru (at the magnetic equator). The top frame shows a 5-minute collection of echoes from 4 heights and 4 frequencies all plotted together. While the 6 frames below it show the skymap from the maximum amplitude echo height after each coherent integration (20 seconds). The east-west separation of echo locations with positive Doppler in the west and negative Doppler in the east suggests eastward propagation of north-south aligned corrugations. This motion continues for several hours. The north-south alignment of the structure suggests that we are seeing field-aligned irregularities which are experiencing a uniform electric field driven (vertical E) drift.

#### Angle of Arrival Spreading for Oblique Incidence

The IONMOD project /14/, an experiment to detect heating effects caused by oblique transmission of high powered HF radio signals provided an opportunity for us to apply Doppler resolution to the task of separating skywave components. A "probe" circuit, a 500 Watt transmitter separated by 2200km from a system of multiple coherent

receiver, was established in the southwestern U.S. If operated at a frequency close to the heater frequency the skywave signal would penetrate the same volume of the E and F-layers that would be modified by the powerful heater signal. Integration periods from 5 to 30 seconds were used and it was found that the elevation angles of arrival of discrete components of nearly equal strength can be determined accurately if the components are resolved in the frequency domain by the Doppler processing. However, when the two components fall in the same Doppler bin the computed angle of arrival is ambiguous.

Figure 4 illustrates the sensitivity of the angle of arrival measurement to the effectiveness of the Doppler resolution technique. The left side of the figure is a "waterfall" display of the middle 400 lines of a 4096 line Doppler spectrum showing the carrier of the received signal. The Doppler range is -6Hz to +6Hz, and the elapsed time between these spectra is 15 seconds. The color of the Doppler bins represents the amplitude of the signal in 4dB steps and a threshold selected by the analyst is applied to eliminate the noise floor. If the amplitude of a Doppler component exceeds this threshold then its phase on the three receivers is used to compute its angle of arrival. The angle of arrival is then plotted on the right side of the display in a color which identifies which part of the Doppler spectrum it came from (ideally we would have 400 colors to uniquely identify each Doppler line in the spectrum). The most prevalent feature observed on this magnetically west-to-east path (and the one shown in Figure 4) was the splitting into high and low rays which would at times provide as much as a few hertz of separation in their Doppler shift. Over the 15 minute period represented in Figure 4 the spectra show the high ray alternately merging and then separating from the low ray such that the plot on the right shows a precise determination of angle of arrival only when the two modes are separated in frequency by more than about 0.1Hz.

There were two major weaknesses to this CW measurement technique which predate the existence of the bistatic oblique propagation mode of the DPS system:

1. With no range resolution, the high and low rays were often overlaying each other when they could have easily been resolved with a modest bandwidth pulsed, chirped or phase coded signal.
2. The amplitude fading due to Faraday rotation, effects was so deep and often so fast that the Doppler spectra showed a typical amplitude modulation spectrum, even though the transmission was an unmodulated CW carrier. A spread spectrum signal would not experience such deep nulls.

## CONCLUSIONS

Resolution of source components by range resolution and spectral analysis combined with interferometer principles is a powerful technique in determining the angle of arrival of superimposed multipath signals.

## REFERENCES

1. Francis, S.H., "Global Propagation of Atmospheric Gravity Waves: A Review," *J. Atmos. Terr. Phys.*, **37**, p. 1011, 1975.
2. Reinisch, B.W., Buchau, J. and Weber, E.J., "Digital Ionosonde Observations of the Polar Cap F Region Convection," *Physica Scripta*, **26**, pp. 372-377, 1987.
3. Buchau, J. and Reinisch, B.W., "Electron Density Structures in the Polar F Region," *Advanced Space Research*, **11**, No. 10, pp. 29-37, 1991.

4. Buchau, J., Weber, E.J., Anderson, D.N., Carlson, Jr., H.C., Moore, J.G., Reinisch, B.W. and Livingston, R.C., "Ionospheric Structures in the Polar Cap: Their Origin and Relation to 250 MHz Scintillation," *Radio Science*, **20**, No. 3, pp. 325-338, May-June 1985.
5. Kelly, M., "The Earth's Ionosphere," Chapter 4, 1989.
6. Reinisch, B.W., "New Techniques in Ground-Based Ionospheric Sounding and Studies," *Radio Science*, **21**, No. 3, May-June 1989.
7. Bibl, K., Reinisch, B.W. and Kitrosser, D.F., "General Description of the Compact Digital Ionospheric Sounder, Digisonde 256," University of Lowell Center for Atmospheric Research, 1981.
8. Golay, M.S., "Complementary Codes," *IRE Transactions on Information Theory*, April 1961.
9. Schmidt, G., Ruster, R. and Czechowsky, P., "Complementary Code and Digital Filtering for Detection of Weak VHF Radar Signals from the Mesosphere," *IEEE Transactions on Geoscience Electronics*, May 1979.
10. Haines, D.M., "Digisonde Portable Sounder (DPS) System Technical Manual," University of Massachusetts Lowell Center for Atmospheric Research, Version 1.2, April 1992.
11. Davies, K., "Ionospheric Radio," *IEEE Series 31*, 1989.
12. Scali, J.L., "Online Digisonde Doppler Analysis," User's Manual, University of Massachusetts Lowell Center for Atmospheric Research, 1992.
13. Reinisch, B.W., Haines, D.M. and Kiklinski, W.S., "The New Portable Digisonde for Vertical and Oblique Sounding," AGARD-CP-502, February 1992.
14. Sales, G.S., Platt, I.G., Huang, Y., Haines, D.M. and Hecksher, J.L., "The Investigation of Changes in the Ionosphere Caused by High Power HF Transmitters," *URSI International Conference on Ionospheric Modification*, September 1991.



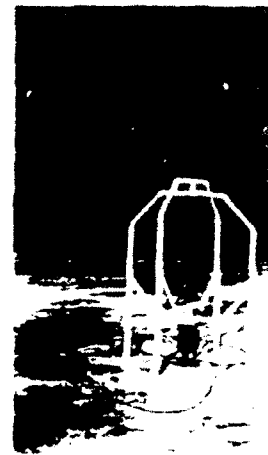
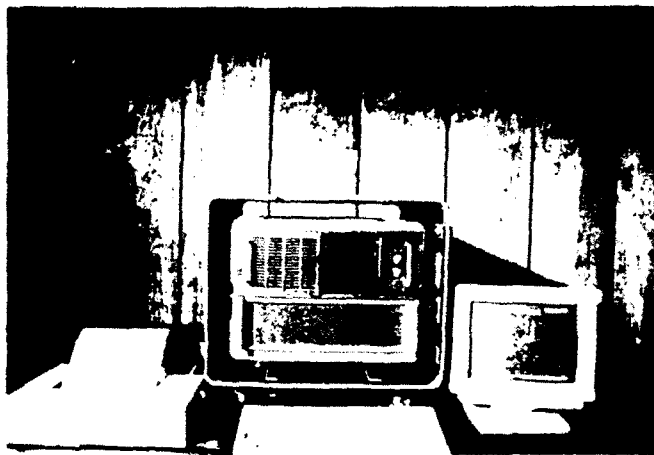


Figure 1 The Digisonde Portable Sounder (DPS) with Display Peripherals and One of its Four Receive Antennas

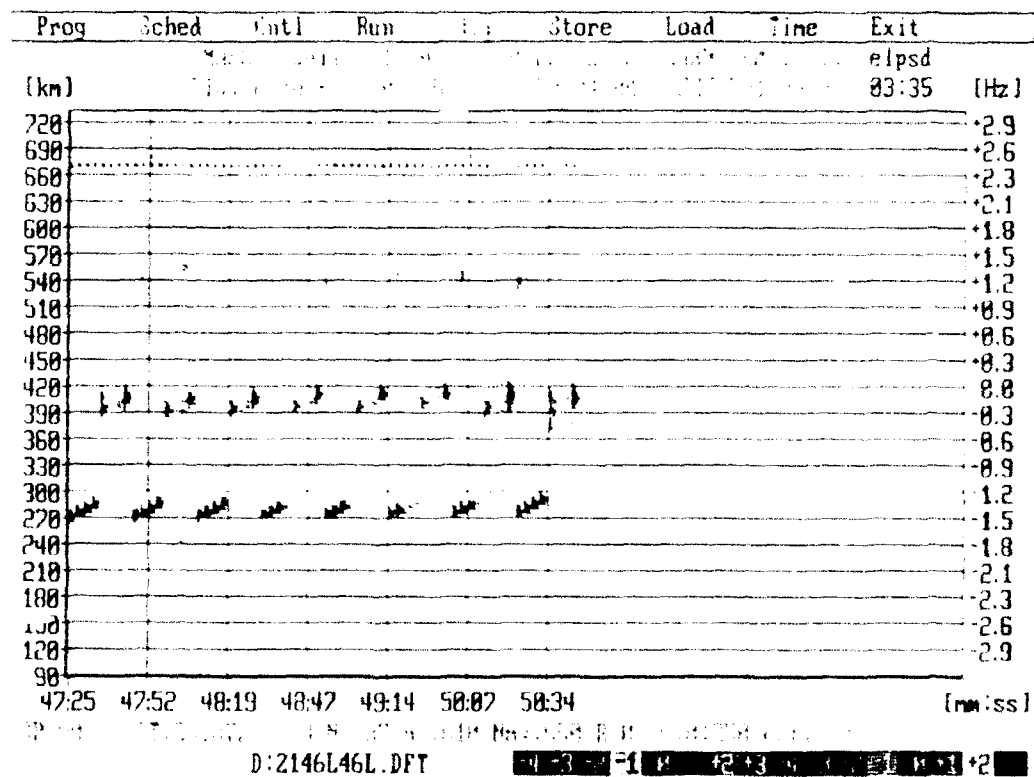


Figure 2 DPS Frequency Multiplexed Doppler Display Showing Four Virtual Height Profiles Followed by the Associated Four Doppler Spectra

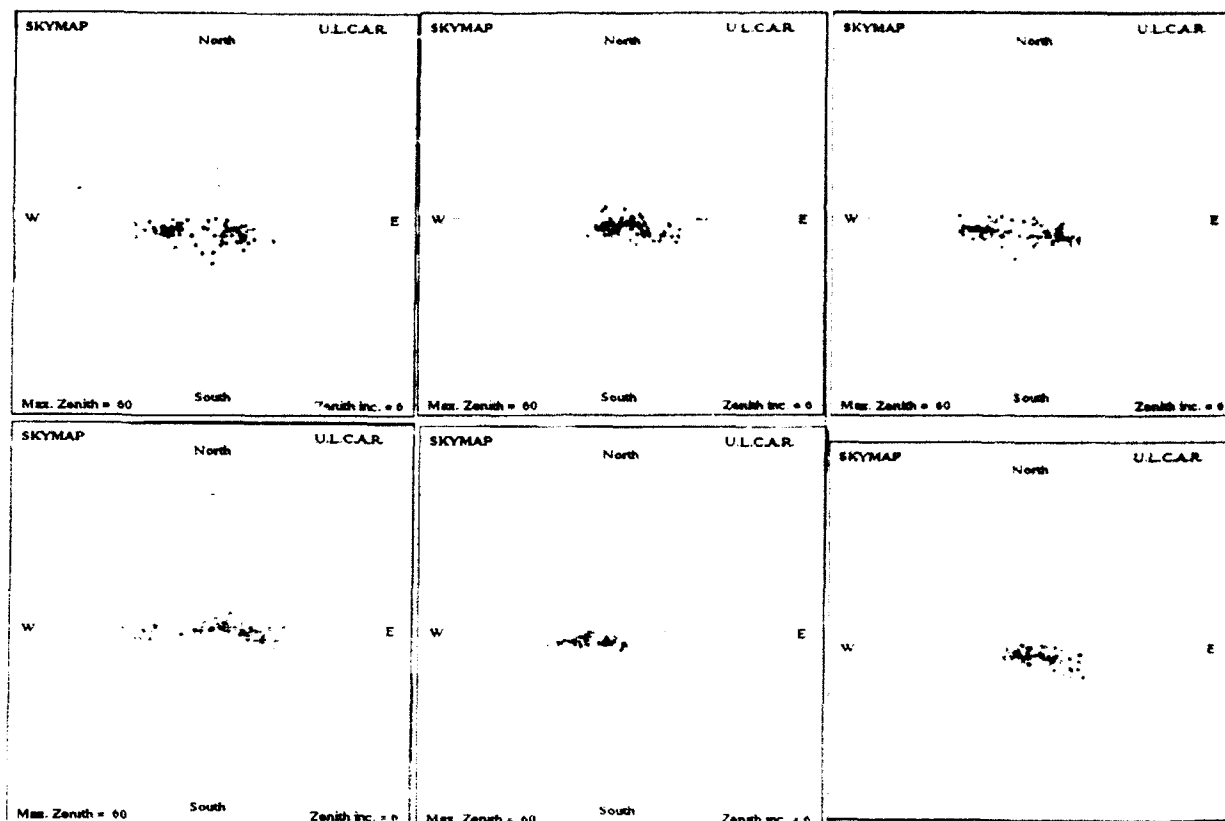


Figure 3. Vertical Incidence Skymap Display Locating Echo Sources Received Over a 20 Second Integration.

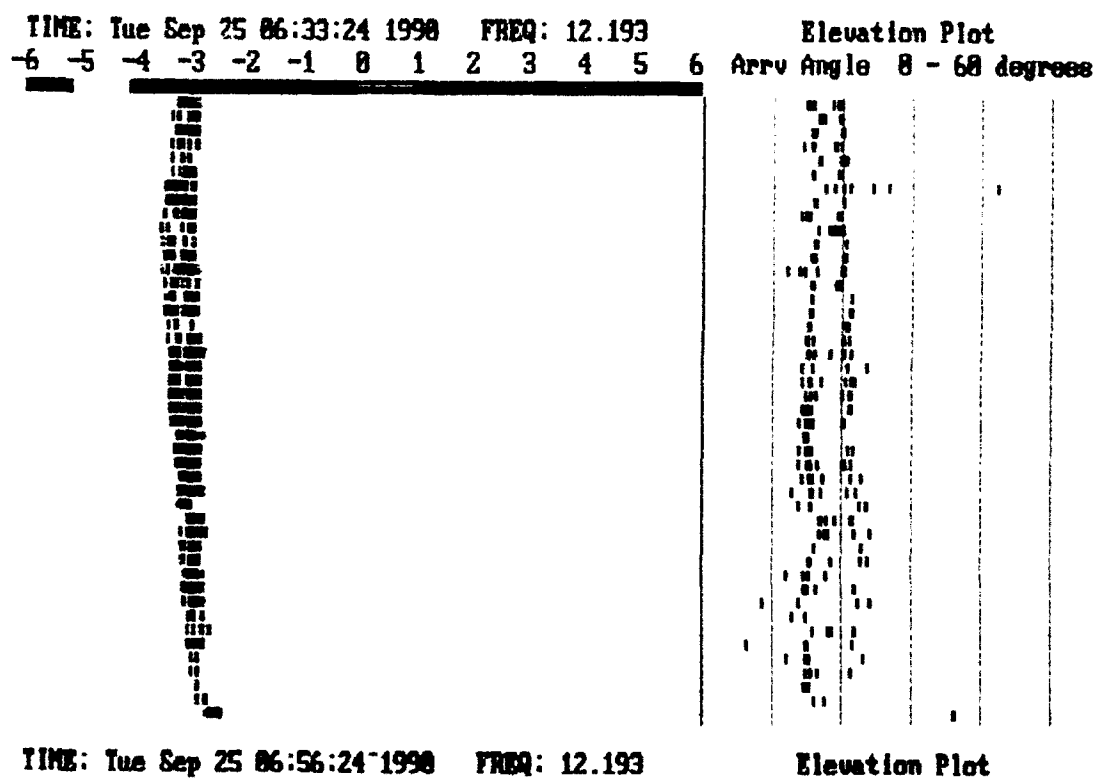


Figure 4. Oblique Signal Angle-of-Arrival After Doppler Domain Resolution 2200km Range.

## DISCUSSION

### M. PITTEWAY

You measure phase as a function of frequency in order to get an accurate group range, so why not measure phase as a function of time in order to get accurate Doppler information, rather than your 128 point FFT to get a Doppler spectrum? (Having heard your verbal answer, to use spectrum to resolve signals with differing Doppler shifts, my (written only) supplementary question is: surely, group range, too, should be invalid to separate different effects; also polarization (implicit in your crossed dipole switching, but what about linear polarization at the equator?) and differing angles of arrival?

### AUTHOR'S REPLY

Getting back to the basic question; we should consider that the FFT process begins by sampling phase (actually complex amplitudes) vs. time. Then rather than trying to measure the phase of the time domain samples we perform a complex FFT which not only separates multiple signal components but provides 21 dB of signal processing gain (improved SNR) making the computed phase of a complex Doppler line significantly more accurate than the original complex amplitude would have been.

Now for the second part. The strength of our technique is that all these discriminating features (group delay, polarization, spatially separated antennas and Doppler shifts) are exploited to ensure that any complex amplitude represents only one sky wave component (i.e., angle of arrival). Perhaps you didn't understand that we build up independent time domain records for each polarization, each antenna, and each frequency (when frequency multiplexing is selected). Then a separate FFT on each record results in a unique Doppler spectrum for each range, polarization, antenna, and frequency; as many as 8192 Doppler spectra (1 million complex Doppler lines) may be computed at the end of one 40 second coherent integration. Such a computation requires about 15 seconds. So the input samples for a DPS Doppler spectrum have already been separated by range, polarization, frequency and antenna before the FFT is performed.

Yes, we used both linear and circular polarization at Jicamarca, Peru (11.5° North magnetic latitude) and found that switching between North-South and East-West magnetic dipole elements (linear polarization) discriminated the O and X waves significantly better than did the circular polarizations.

### C. GOUTELARD

Commentaire:

Vous dites ne pas pouvoir séparer les modes ordinaires et extraordinaires avec une analyse de Fourier. Le rapport signal/bruit que vous obtenez après traitement vous permettrait d'utiliser des méthodes haute résolution telles que MUSIC. Nous appliquons cette méthode et nous mettons clairement en évidence les modes ordinaires et extraordinaires.

Question:

Quelles sont les dimensions minimales des irrégularités que vous pouvez observer et avez-vous tenté d'en construire des modèles?

Comment:

*You say that you cannot separate the ordinary modes from the extraordinary modes using a Fourier analysis. With the signal to noise ratios you obtain after processing you could use high resolution methods such as MUSIC. We apply this method and are able to clearly distinguish ordinary from extraordinary modes.*

Question:

What are the minimum sizes of the irregularities which you are able to observe and have you attempted to construct models?

### AUTHOR'S REPLY

After a 20 second integration of 4 frequencies, 4 antennas and 2 polarizations at each of 256 sampled heights (range bins) we have 8192 FFTs to perform before the next measurement can be made. A 10 MIP 16-bit fixed point signal processor can accomplish this in about 5 seconds, while a MUSIC implementation would require a floating point processor and would take much longer to analyze the 8192 spectral points. The data shown was a received CW transmission (no range resolution). If the experiment is done again using the DPS system which provides 10 km range resolution, the o and x components would be resolved most of the time. The sky map resolution, or accuracy is about 1° which translates to 3 km of F-region heights. We find a continuum of scatterer sources over these distance but it is uncertain whether this spreading is apparent, due to imperfect signal-to-noise ratio, or is a real geophysical phenomenon. Structures separated by tens of kilometers are easily seen and the polar convection plasma patches are typically separated by hundreds of kilometers as they pass overhead.

### G. HAGN

Autoscaling is still somewhat of an art, so ARTIST is an appropriate acronym. You showed a beautiful example of a mid-latitude ionogram which was autoscaled. You also showed a rather complicated ionogram from a high latitude (with 3 different electron densities), and it seemed that the autoscaler had difficulty with this ionogram. Can you comment on this and on the state-of-the-art of autoscaling?

### AUTHOR'S REPLY

This ionogram had at least three different F-layer traces due to plasma patches overhead which had different electron density profiles. The automatically scaled F-layer trace jumped off the correct trace and continued on one of the obliquely received traces. This highlighted the need to distinguish between overhead (e.g.  $\pm 22^\circ$  zenith angle) echoes and oblique which we are now adding to the software using a simple digital beamforming algorithm. The older Digisonde 256 successfully used analog beamforming to accomplish this discrimination, and the ARTIST uses this information.

# UTILISATION DU RESEAU DE RECEPTION DU SONDEUR A RETRODIFFUSION DE L'ILE LOSQUET EN RADIOGONIOMETRIE

J.Y. LE SAOUT, F. GAUTHIER, N. RUELLE, R. FLEURY

Centre National d'Etudes des Télécommunications  
LAB/PTI/GER, Route de trégastel  
B.P. 40, 22301 Lannion Cedex, France

## RESUME

On décrit la méthode de radiogoniométrie à station unique mise en oeuvre à partir du sondeur à rétrodiffusion de l'île Losquet. Des résultats expérimentaux de détermination de l'azimut et de l'élévation d'ondes ayant transitées par l'ionosphère, obtenus à partir d'émetteurs bien localisés, sont présentés. Les angles mesurés sont comparés, à des résultats issus d'un modèle de propagation et à des directions d'arrivée calculées à partir d'une transposition en oblique d'ionogrammes verticaux. On montre que:

- l'azimut dans lequel se trouve l'émetteur peut être déterminé avec une précision qui est celle du pas de mesure,
- en élévation, la méthode permet la mesure des angles associés aux trajets dominants bien que ses performances soient limitées en présence de multi-trajets.

## 1 INTRODUCTION

Ces dernières années de nombreuses méthodes de radiogoniométrie ont été publiées [1], [2], [3]. Ces différentes techniques peuvent être classées en trois grandes catégories :

- les méthodes interférométriques,
- les méthodes à balayage Doppler,
- les méthodes à balayage d'antenne.

La radiolocalisation est l'un des objectifs de ces études [4].

Dans le cadre d'études de propagation HF en ondes de ciel, la radiogoniométrie représente un moyen d'investigation du milieu ionosphérique. Lorsque la position de l'émetteur est connue la détermination d'angles d'arrivée d'ondes ayant transitées par l'ionosphère nous renseigne sur l'altitude virtuelle du point de réflexion de ces ondes. Ces informations sont complémentaires de celles obtenues par sondage oblique.

L'objectif de cet article est de présenter des résultats de radiogoniométrie à station unique par une technique de balayage d'antenne, obtenus à l'aide du réseau de réception du sondeur à rétrodiffusion du Centre National d'Etudes des Télécommunications (CNET).

On décrit la méthode employée et on présente des résultats de mesures réalisés à différentes fréquences et pour plusieurs distances, à partir d'émetteurs de radiodiffusion bien localisés. Ces

résultats concernent à la fois la détermination de l'azimut et de l'élévation des ondes reçues.

Après quoi on compare ces angles déterminés expérimentalement, à ceux issus d'un modèle de prévisions ainsi qu'aux angles calculés par transposition en oblique d'ionogrammes verticaux.

Enfin on discute des limitations de la méthode.

## 2 PRESENTATION DE LA METHODE

### 2.1 Le système d'aériens

Le sondeur à rétrodiffusion du CNET décrit par ailleurs [5] est implanté en Bretagne (France), sur l'île Losquet (48.8° N ; 3.6° W). Ce système utilise pour l'analyse de l'ionosphère les fréquences radioélectriques comprises entre 6 et 30 MHz. Son système d'aériens est constitué de deux réseaux circulaires concentriques d'antennes : un pour l'émission, l'autre pour la réception.

La méthode de radiogoniométrie présentée dans cet article met uniquement en oeuvre le réseau de réception du sondeur.

Ce réseau est constitué (Figure 1):

- de 64 antennes actives, à polarisation Verticale, réparties sur un cercle de 68.8 mètres de rayon,
- de 64 déphaseurs à commande numérique associés individuellement à chacune des antennes. Ces derniers permettent un contrôle continu du déphasage des signaux reçus de 0° à 360°,
- d'un récepteur de mesure,
- d'un calculateur qui permet la commande et la gestion informatique du récepteur et des déphaseurs ainsi que l'acquisition et l'enregistrement des mesures.

### 2.2 Principe de la méthode

Dans un réseau de N antennes, le signal  $S_i$  disponible à la sortie d'une antenne  $i$  est pondéré par un coefficient complexe  $W_i$ . Les signaux  $W_i \times S_i$  sont alors sommés avant l'entrée du récepteur.

$$S_r = \sum_{i=1}^N W_i \cdot S_i$$

La conformation du *diagramme de directivité* consiste à déterminer les coefficients  $W_i$  de manière à maximiser l'énergie dans la direction de visée. Les pondérations  $W_i$  utilisées sont des termes de phase appliqués aux signaux  $S_i$  par les déphaseurs à commande numérique. Ils s'expriment de la façon suivante :

$$W_i(Az_0, E_0) = \exp \left\{ -j \frac{2\pi R}{\lambda} \cdot \cos E_0 \cdot \cos(\varphi_i - Az_0) \right\}$$

où :

$Az_0$  et  $E_0$  sont respectivement les angles de pointage du lobe principal en azimut et en élévation (ils définissent la direction de visée).

$\lambda$  est la longueur d'onde

$R$  est le rayon d'implantation des antennes

$\varphi_i$  est l'angle au centre du réseau de l'antenne  $i$  par rapport au nord géographique, compté positivement vers l'Est.

La figure 2a présente un exemple de simulation du diagramme de directivité du réseau de réception obtenue à la fréquence de 20 MHz et pour des angles de visée de 90° en azimut et de 30° en site. La mesure de ce diagramme effectuée à l'aide d'une technique décrite par ailleurs [6] est présentée figure 2b. Le bon accord entre la simulation et la mesure confirme l'aptitude des déphaseurs à conformer le diagramme de directivité dans la direction choisie.

Le récepteur réalise une mesure du module de  $S_T$  pour chaque direction de visée et la représentation de  $|S_T|$  en fonction de la direction de visée est appelée *diagramme de réception*.

Dans l'hypothèse d'une onde plane incidente le diagramme de réception présente un maximum dans la direction d'arrivée de l'onde et des lobes secondaires d'amplitude moindre (environ -8 dB) dans d'autres directions.

La technique de radiogoniométrie que nous appelons *balayage de lobe* est un balayage de l'espace par le diagramme de directivité du réseau de réception.

En pratique les deux angles d'arrivée de l'onde sont généralement inconnus. Expérimentalement on détermine dans un premier temps l'azimut en se donnant une élévation a priori. L'azimut étant alors connu, on effectue un balayage en site.

### 3 RESULTATS

#### 3.1 Résultats expérimentaux

Nous avons expérimenté la méthode présentée ci-dessus sur des émetteurs de radiodiffusion car la connaissance de leur coordonnées géographiques est une aide à la validation de nos mesures. La distance entre ces émetteurs et l'île Losquet est comprise entre 500 et 1500 km environ.

Les résultats que nous présentons ont été obtenus en utilisant la technique du balayage de lobe, en azimut de 0° à 360° et en élévation de 0° à 90°. Pour tenir compte des variations d'amplitude observées

(modulation d'amplitude et évanouissement associés à la présence de multi-trajets) dix balayages successifs ont été moyennés [7]. La durée d'un balayage dépend :

- (i) du temps nécessaire à l'adressage des déphaseurs et à la réalisation de la mesure d'amplitude par le récepteur (environ 0,45 seconde).
- (ii) du pas angulaire de mesure qui est choisi par l'expérimentateur en fonction des conditions de propagation et de la fréquence de l'émetteur observé.

La figure 3 présente un exemple de radiogoniométrie effectuée à partir de l'émetteur de Lenk (46°27' N, 7°27' E) à la fréquence de 13,685 MHz. Le diagramme azimutal (figure 3a) a été obtenu en faisant varier l'angle de pointage du lobe principal en azimut par pas de deux degrés, ce qui correspond à une durée d'acquisition totale d'environ 14 minutes. Le diagramme en élévation (figure 3b) a été obtenu en faisant varier l'angle en site par pas de 2 degrés et conduit à une durée d'acquisition d'environ 3 minutes. Les figures 3a et 3b montrent clairement une direction privilégiée d'arrivée des signaux, traduisant la présence d'un trajet dominant de propagation. Par ailleurs, l'azimut dans lequel sont observés les signaux ( $103^\circ \pm 1^\circ$ ) est le même que celui calculé géographiquement ( $103^\circ$ ).

La figure 4 est un autre exemple représentatif d'un mode de propagation dominant, qui a été obtenu sur l'émetteur de Wertachtal (48°7' N, 11°5' E) à la fréquence de 17,845 MHz. Les durées d'acquisition en azimut et en élévation sont environ de 13 et 7 minutes respectivement. La comparaison des figures 3a et 4a (obtenues par pas de 2 degrés) montre que l'ouverture du lobe principal en azimut décroît avec la longueur d'onde. L'incertitude associée à la mesure des directions d'arrivée des signaux en azimut varie dans le même sens. Par ailleurs, l'azimut dans lequel sont observés les signaux ( $89^\circ \pm 1^\circ$ ) est le même que celui calculé géographiquement ( $89^\circ$ ). Ce résultat est représentatif de l'ensemble des mesures que nous avons effectuées : le tableau 1 présente les angles de gisement mesurés et calculés à partir des positions géographiques des émetteurs. La comparaison des angles indique que pour les géométries présentées, la dispersion azimutale introduite par l'ionosphère est inférieure au pas de mesure.

La comparaison des figures 3a et 4a avec les figures 3b et 4b montre que le secteur angulaire des directions d'arrivée observées est plus important en élévation qu'en azimut. Ce résultat peut être expliqué :

- (i) par la valeur de l'ouverture du lobe principal plus importante en élévation qu'en azimut. Le tableau 2 indique la valeur des ouvertures à 3 dB calculées à l'aide de simulations pour des fréquences comprises entre 6 et 30 MHz et des élévations de visée de 20°, 30° et 40°.
- (ii) par la présence des deux modes de polarisation ordinaire et extraordinaire par trajet de propagation.
- (iii) par une dispersion angulaire d'origine ionosphérique plus élevée en site qu'en gisement. Celle-ci est d'autant plus importante qu'il existe des gradients.

Il est important de remarquer que les mesures que nous présentons sont le résultat de convolutions entre les diagrammes de directivité et des "plans d'ondes" incidents. La séparation entre les effets de

la largeur du lobe et de la dispersion angulaire introduite par l'ionosphère sur les mesures en site n'est donc pas immédiate.

Si les figures 3b et 4b sont représentatives de la présence d'un mode de propagation dominant, dans le cas général la stratification verticale de l'ionosphère est à l'origine de trajets multiples : la figure 5 est un résultat de mesure mettant en évidence ce type de situation. Le diagramme de réception en site possède deux lobes principaux indiquant a priori la présence de deux trajets distincts en élévation.

Il est à noter que comparée à la mesure d'un seul trajet la présence de plusieurs chemins de propagation se traduit par une déformation du diagramme de réception qui évolue au cours du temps en fonction des relations de phase entre les aériens. Suivant l'instant de la mesure, la position et le nombre des lobes principaux peut varier et rendre ainsi nécessaire le moyennage de plusieurs balayages consécutifs.

Dans le cas de la figure 5 l'écart angulaire entre les deux trajets est supérieur à l'ouverture du lobe principal. La figure 6 présente un résultat de balayage obtenu en présence de multi-trajets pour lequel la différence d'élévation entre les chemins de propagation est comparable à l'ouverture du lobe principal. La discrimination devient alors difficile.

Les résultats que nous avons présentés dans ce paragraphe montrent que la détermination de l'azimut pose peu de difficulté. Par contre, en site la présence de multi-trajets peut engendrer des problèmes d'interprétation. Nous nous proposons d'aborder leur analyse dans le paragraphe suivant

### 3.2 Interprétation des mesures en site.

L'évaluation des performances de la technique de radiogoniométrie, repose sur deux approches :

- La première consiste à comparer les angles de site mesurés aux directions d'arrivée calculées à l'aide du modèle de prévisions de propagation ASAPS [8]. Ce modèle permet la caractérisation des modes de propagation et le calcul des angles d'incidence qui leur sont associés, pour une fréquence, une liaison et une date données. Il est à noter que ce modèle ne fournit des informations que sur l'état moyen de l'ionosphère.

- La seconde méthode est une comparaison de ces mêmes mesures aux angles issus de la transposition en oblique des ionogrammes verticaux obtenus à Lannion. Cette transformation, effectuée à l'aide de la loi de la sécante [9], permet le calcul des angles d'incidence pour une fréquence et une distance de propagation données. Rappelons que cette méthode ne s'applique qu'aux modes ordinaires et ne prend pas en compte d'éventuels gradients horizontaux de densité électronique.

Par ces deux méthodes, la qualité des mesures a été approchée en comparant la différence entre les angles mesurés et calculés avec le pas de mesure. Pour simplifier l'analyse des résultats seuls les angles associés au maximum du diagramme de réception ont été utilisés. Nous avons par ailleurs séparé les mesures en deux catégories correspondant aux bandes de fréquences 6-12 MHz et 12-22 MHz; ce qui permet de prendre en compte dans l'analyse des résultats, le fait que le diagramme de directivité possède un lobe principal plus large aux basses fréquences.

Les trois premières colonnes du tableau 3 résument les principales caractéristiques (lieu de l'émetteur, fréquence utilisée, date, heure) des mesures effectuées au cours du mois d'Octobre 1991 et du mois de Janvier 1992, à partir d'émetteurs de radiodiffusion dans la bande 12-22 MHz. Les angles de site déterminés et le pas de mesure utilisé sont représentés dans les colonnes 4 et 5. Les valeurs soulignées correspondent au maximum d'amplitude observé. Dans les quatre dernières colonnes du tableau 3 figurent les angles d'incidences obtenus à l'aide du modèle ASAPS (méthode 1) et les directions d'arrivée issues des transpositions vertical-oblique (méthode 2), ainsi que les différences respectives de ces angles par rapport aux mesures.

On constate que :

- lorsque une ou deux directions d'arrivée sont mesurées, l'angle associé au maximum d'amplitude observé correspond à un des trajets prévus par les méthodes 1 et 2 : l'écart entre les angles calculés et mesurés est de l'ordre de grandeur du pas de mesure.

- lorsque le nombre de trajets mesurés est supérieur à deux, les écarts entre les angles mesurés et ceux issus du calcul sont plus importants. Ce résultat semble indiquer que l'incertitude sur la mesure augmente avec le nombre de trajets mesurés.

Les résultats de radiogoniométrie obtenus dans la bande des fréquences 6-12 MHz (Tableau 4) indiquent que les différences entre les mesures et les prévisions sont plus élevées que celles observées dans la bande 12-22 MHz. L'ouverture du lobe principal du diagramme de directivité, plus important aux basses fréquences qu'aux fréquences élevées est probablement à l'origine de ce résultat. Cependant un nombre plus important de mesures est nécessaire pour confirmer cette explication.

## 4 LES LIMITES DE LA METHODE

Les limites de la technique du balayage de lobe ont été évaluées en ayant recouru à des simulations utilisant un modèle de signal développé par Bertel et al [10], [11]. Ce modèle prend en compte les trajets multiples, la polarisation des modes propagés et le décalage Doppler différentiel existant entre ces derniers. On présente ici des résultats relatifs à l'étude de deux trajets.

### 4.1 L'ouverture du lobe principal.

Nous avons effectué des simulations en utilisant deux trajets séparés d'un écart angulaire inférieur à la largeur à 3 dB du lobe principal du diagramme de directivité. La figure 7 montre que dans ce cas le diagramme de réception présente un élargissement du lobe principal. La valeur de l'angle correspondant au maximum du diagramme (28°) est la même que l'élévation d'un des trajets introduits. Il est à noter que de manière générale l'angle associé au maximum est intermédiaire entre les deux élévations d'arrivée.

Lorsque l'écart angulaire et l'ouverture à 3 dB du lobe principal sont dans un rapport 2, le diagramme de réception permet une discrimination correcte des directions d'arrivée.

La limitation introduite par la largeur du lobe principal diminue avec la longueur d'onde pour un même écart angulaire, puisque l'ouverture décroît lorsque la fréquence augmente.

#### 4.2 Le rapport des amplitudes.

Le niveau des lobes principaux du diagramme de réception est fonction de l'amplitude relative des signaux à la sortie de l'ionosphère et du diagramme de rayonnement des antennes. Dans ces conditions, les trajets de faible amplitude peuvent être masqués par les lobes secondaires associés au mode dominant.

Nous avons simulé ce phénomène en faisant varier les amplitudes associées à deux trajets d'incidences  $15^\circ$  et  $35^\circ$ . La figure 8 montre qu'avec des amplitudes introduites identiques, la méthode permet de retrouver les angles d'arrivée. Par contre, lorsque les énergies sont dans un rapport 0.5, le trajet de plus faible amplitude peut être confondu avec un des lobes secondaires du diagramme de réception: ceci est illustré par la figure 9.

#### 4.3 La durée de la mesure.

Rappelons qu'un balayage en site de  $0$  à  $90^\circ$  par pas de  $2^\circ$  dure environ 20 secondes. La non stationnarité du milieu pendant cette durée, lors de la présence de multi-trajets, est aussi à l'origine d'une déformation du diagramme de réception. Celle-ci est d'autant plus importante que la période des variations temporelles des Doppler différentiels est de l'ordre de grandeur de la durée du balayage. Dans ces conditions le diagramme de réception présente un "feuilletage" illustré par la figure 10.

De manière générale, les résultats de simulations montrent que diverses influences interviennent dans la déformation du diagramme de réception. Cependant les maximums observés permettent une interprétation correcte de la simulation avec deux trajets.

### 5 CONCLUSION

Les résultats de l'interprétation des mesures de radiogoniométrie que nous avons présentés dans cet article peuvent être résumés comme suit :

En azimut, la comparaison des directions d'arrivée mesurées avec celles calculées à partir des positions géographiques des émetteurs montre que l'azimut dans lequel se trouve l'émetteur peut être déterminé avec une précision qui est celle du pas de mesure utilisé.

En site, la comparaison des angles mesurés avec ceux issus du modèle de prévision ASAPS ainsi que ceux qui ont été calculés par transposition oblique d'ionogrammes verticaux montre que : lorsque un ou deux trajets sont observés, la différence entre l'angle associé au maximum du diagramme de réception et ceux qui ont été calculés est de l'ordre de grandeur du pas de mesure. Cette différence est plus importante aux basses fréquences.

Par ailleurs, les résultats de la simulation des mesures en site indiquent que les performances de la méthode dépendent du nombre de trajets présents au moment de la mesure : la détermination de l'angle d'incidence associé à un seul trajet peut être effectuée sans difficulté. Par contre, la discrimination de multi-trajets n'est possible que si les amplitudes relatives qui leur sont associées sont dans un rapport au moins de 0.5 et si l'écart angulaire entre les directions d'arrivée est supérieur à l'ouverture à 3 dB du lobe principal du diagramme de directivité.

Les résultats de cette étude préliminaire montrent que l'on peut envisager l'emploi de la technique du balayage de lobe pour des études de radiogoniométrie HF telles que la radiolocalisation à station unique ou la mesure en temps réel des conditions de propagation.

### REMERCIEMENTS

Les auteurs remercient l'ensemble de l'équipe "sondeur à rétrodiffusion", pour leur contribution efficace à la partie expérimentale de cette étude.

### REFERENCES

1. Gething, P.J.D., "Radio Direction Finding", IEE Electromagnetic waves series 4, Published by Peter Peregrinus Ltd, 1978.
2. Multedo, G., "Les techniques d'interférométrie utilisées dans les radiogoniomètres à THOMSON-CSF", Revue technique Thomson-CSF, juin 1987, vol 19, N°2.
3. McNamara, L.F., "Ionospheric modeling in support of single-station-location of long range transmitters", J. Atmos. Terr. Phys. 50, 1988, pp781-795.
4. Fombonne, P., "Radionavigation Radiolocalisation", Masson 1983.
5. Le Saout, J.Y., Bertel, L., "Antenna arrays of the C.N.E.T. backscatter HF radar", Fourth International Conference on HF radio systems and techniques, april 1988, IEE Conference Publication N°284, pp 280-284.
6. Le Saout, J.Y., Bertrand, P., Poitevin, R., "Les caractéristiques expérimentales des systèmes d'aériens du radar HF à rétrodiffusion", Journées Internationales de Nice sur les Antennes, novembre 1990, pp 572-575
7. Bertel, L., Gourvez, P., Le Saout, J.Y., Ruelle, N., "HF Receiving antenna array behaviour", Sixth International Conference on Antennas and Propagation 1989, Conference Publication N°301, Part 1, pp 507-511.
8. ASAPS, Advanced Stand Alone Prediction System, User guide, IPS Radio and Space Service, Australia, 1989.
9. Davies, K., "Ionospheric Radio" IEE Electromagnetic waves series 31, Published by Peter Peregrinus Ltd, 1990.
10. Bertel, L., Gourvez, P., Leclerc, P., Le Saout, J.Y., Rojas Valera, J., "Etude du comportement d'un réseau d'antennes HF en réception", Journées Internationales de Nice sur les Antennes, novembre 1986.
11. Bertel, L., Le Ray, B., Baltazart, V., Le Saout, J.Y., "Analysis of behaviour of small basis radiogoniometers using interferometry technique", Conference on solar terrestrial and space physics and workshop on ionospheric modelling an radio wave propagation, Melbourne, february 1990.

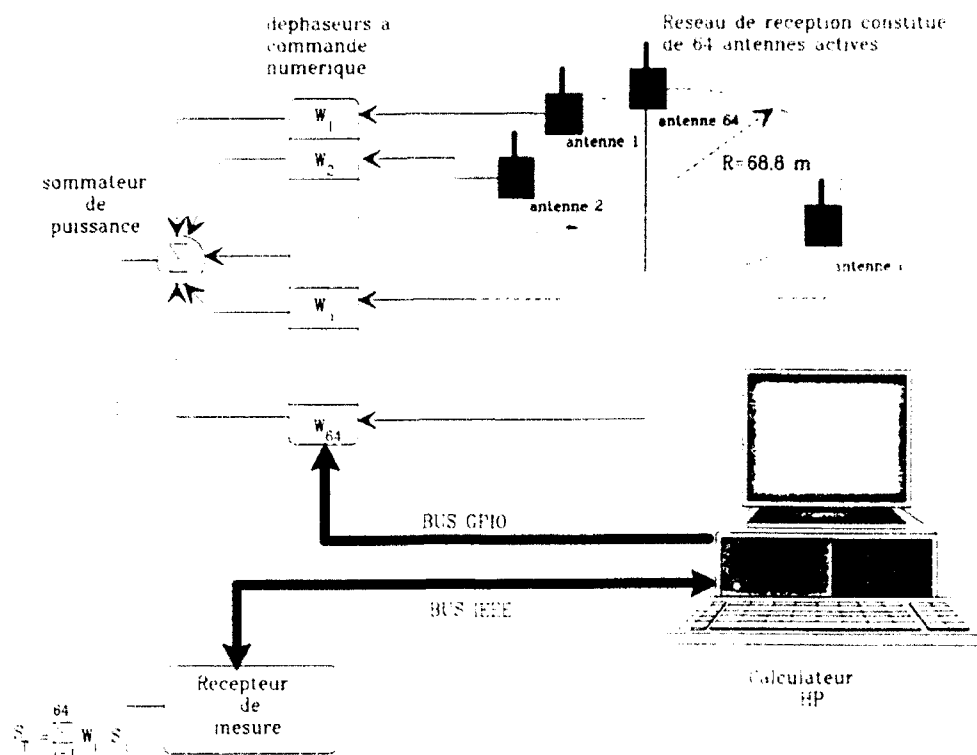


Figure 1 : Synoptique du réseau de réception du sondeur à rétrodiffusion du CNET lors de son utilisation en radiogoniométrie.

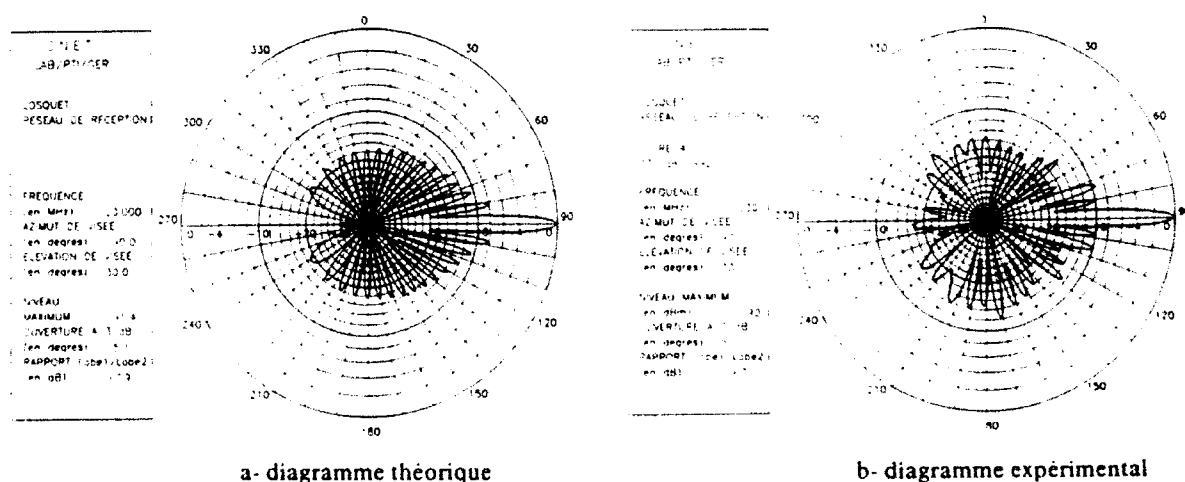
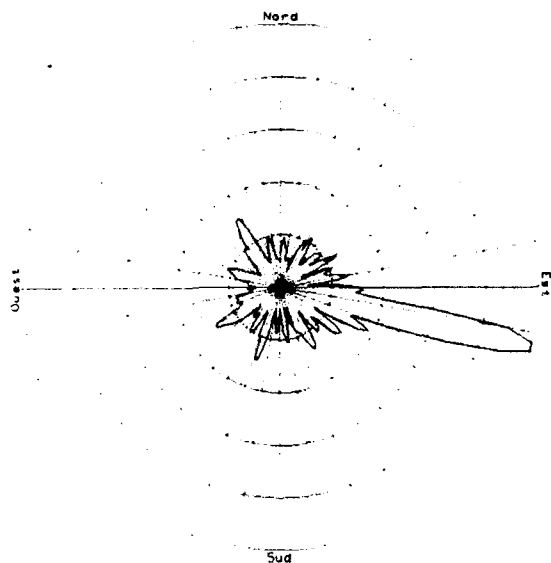


Figure 2 : Diagrammes de directivité théorique et mesuré du réseau de réception à la fréquence de 20 MHz.



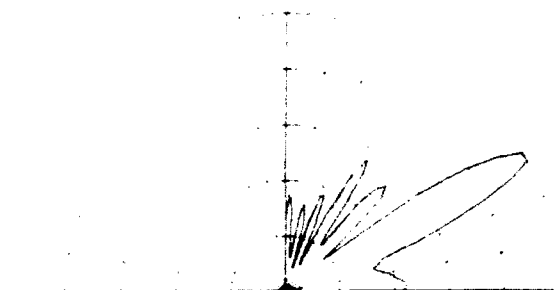
FICHIER MGISE81	AZI.MINI : 0°	PAS: 2°
TYPE DE VISU.: LIN.	AZI.MAXI : 360°	MAXI: -28.66 dBm
DATE : 21 Oct 1991	FREQUENCE : 13.685MHz	MOY. CALCULEE
HEURE : 09:05:53 T.U.	ELEVATION : 20°	SUR 10 Fichiers



a- diagramme azimuthal

FICHIER MSITE32	ELEV.MINI : 0°	PAS : 2°
TYPE DE VISU.: LIN.	ELEV.MAXI : 90°	MAXI: -27.87 dBm
DATE : 21 Oct 1991	FREQUENCE : 13.685MHz	MOY. CALCULEE
HEURE : 09:11:16 T.U.	AZIMUT : 100°	SUR 10 Fichiers

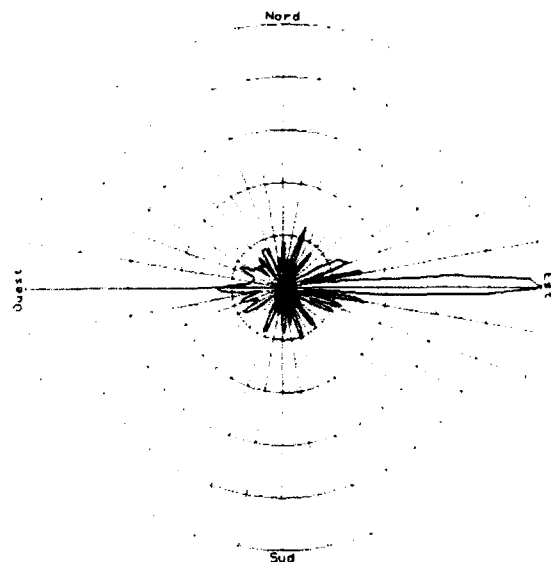
COORDONNEES POLAIRES EN ELEVATION



b- diagramme en site

Figure 3 : Diagrammes de réception obtenus à partir de l'émetteur de radiodiffusion de Lenk (Suisse)

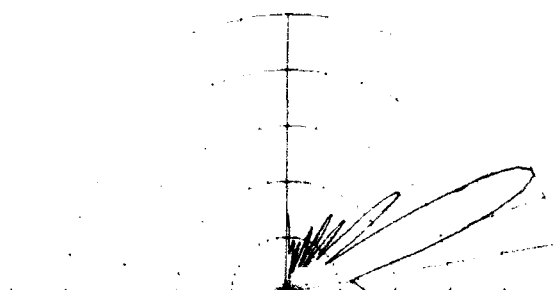
FICHIER MGISE22	AZI.MINI : 0°	PAS: 2°
TYPE DE VISU.: LIN.	AZI.MAXI : 360°	MAXI: -38.87 dBm
DATE : 21 Oct 1991	FREQUENCE : 17.845MHz	MOY. CALCULEE
HEURE : 13:59:06 T.U.	ELEVATION : 26°	SUR 10 Fichiers



a- diagramme azimuthal

FICHIER MSITE2	ELEV.MINI : 0°	PAS : 1°
TYPE DE VISU.: LIN.	ELEV.MAXI : 90°	MAXI: -41.34 dBm
DATE : 21 Oct 1991	FREQUENCE : 17.845MHz	MOY. CALCULEE
HEURE : 14:08:01 T.U.	AZIMUT : 90°	SUR 10 Fichiers

COORDONNEES POLAIRES EN ELEVATION



b- diagramme en site

Figure 4 : Diagrammes de réception obtenus à partir de l'émetteur de radiodiffusion de Wertachtal (Allemagne).

TX	FREQUENCE (en MHz)	DATE HEURE	ELEVATION DE VISEE	LOSQUET	PAS	AZIMUT GEOMETRIQUE
WIEN	6.155	08/10/91 10:04	50°	83°	5°	86°
	15.45	09/10/91 09:25	15°	86°	2°	86°
	21.49	21/10/91 14:26	16°	86°	2°	86°
LENK	9.535	22/10/91 15:35	23°	103°	2°	103°
	13.685	21/10/91 09:05	28°	104°	2°	103°
ALLO	6.17	08/10/91 10:14	50°	112°	5°	114°
SKEL	9.76	22/10/92 14:25	35°	4°	3°	5°
ARGA	11.79	09/10/91 08:46	25°	178°	3°	176°
JUNG	6.09	08/10/91 09:32	20°	76°	5°	79°
WERT	17.845	21/10/91 13:59	26°	89°	2°	89°

Tableau 1 : Comparaison des angles azimutaux mesurés, avec ceux calculés à partir des positions géographiques des émetteurs de:

WIEN (48°00 N, 16°28 E)  
 LENK (46°27 N, 07° 27 E)  
 ALLOUIS (47° 00 N, 02°00 E)  
 SKELTON (54° 44 N, 02°54 W)  
 ARGANDA (40° 18 N, 03° 31 W)  
 JUNGLINSTER (49°40 N, 06°19E)  
 WERTACHTAL (48°05N, 10°41E)

FREQUENCE (en MHz)	OUVERTURE EN AZIMUT: $E_0=30^\circ$ (en deg.)	OUVERTURE EN SITE: $E_0=20^\circ$ (en deg.)	OUVERTURE EN SITE: $E_0=30^\circ$ (en deg.)	OUVERTURE EN SITE: $E_0=40^\circ$ (en deg.)
6	17.3	28.2	33.6	25.3
10	10.4	23.6	19.7	14.3
14	7.4	20.0	13.4	10.1
18	5.7	16.5	10.3	7.8
22	4.7	13.1	8.3	6.4
26	4.0	10.8	7.0	5.4
30	3.4	9.2	6.0	4.6

Tableau 2: Ouverture théorique à 3 dB du lobe principal du diagramme de directivité.

FICHER MSITE6	ELEV.MINI : 0°	PAS : 2°
TYPE DE VISU.: LIN.	ELEV.MAXI : 90°	MAXI: -31.74 dBm
DATE : 17 Jan 1992	FREQUENCE : 9.76Mhz	MOY. CALCULEE
HEURE : 09:52:27 T.U.	AZIMUT : 5°	SUR 18 Fichiers

COORDONNEES POLAIRES EN ELEVATION

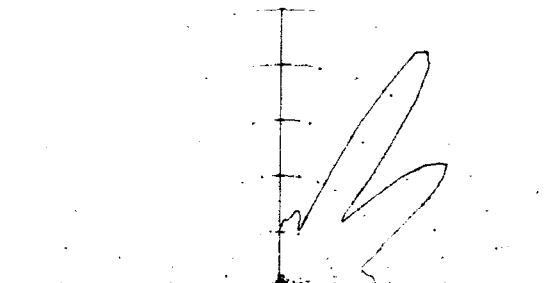


Figure 5: Diagramme de réception en site obtenus à partir de l'émetteur de radiodiffusion de Skelton.

FICHER MSITE91	ELEV.MINI : 0°	PAS : 2°
TYPE DE VISU.: LIN.	ELEV.MAXI : 90°	MAXI: -41.14 dBm
DATE : 18 Oct 1991	FREQUENCE : 13.73Mhz	MOY. CALCULEE
HEURE : 09:37:43 T.U.	AZIMUT : 86°	SUR 18 Fichiers

COORDONNEES POLAIRES EN ELEVATION

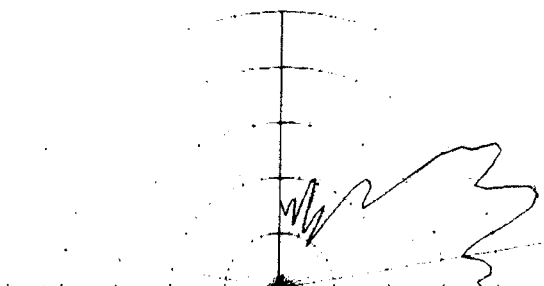


Figure 6: Diagramme de réception en site obtenu à partir de l'émetteur de radiodiffusion de Lenk.

TX	FREQUENCE (en MHz)	DATE HEURE (T.U.)	SITES LOSQUET en deg.	PAS en deg.	SITES METHODE 1 en deg.	ECART en deg.	SITES METHODE 2 en deg.	ECART en deg.
WIEN	13.73	18/10/91 09:37	5 49 <u>12</u> 31	2	7 (1E) 18 (1F) 37 (2F)	-1	8.0 (1E) 48.2 (3F2) 14.0 (1F1) 33.2 (2F2)	-5
	15.45	18/10/91 09:23	7 31 47	1	17 (1F) 38 (2F)	10	15.5 (1F2) 34.4 (2F2)	8.5
	15.45	17/01/92 09:35	<u>13</u> 47	2	16 (1F) 35 (2F)	-1	13.7 (1F2) 34.8 (2F2)	-1.3
	21.49	21/10/91 10:00	<u>17</u>	1	18 (1F)	1	5.1 (1Es) 16.2 (1F2)	-0.8
	21.49	17/01/92 09:50	<u>16</u> 40	2	16 (1F)	0	16.0 (1F2)	0
LENK	13.685	18/10/91 10:05	<u>28</u> 45	2	30 (1F) 53 (2F)	2	13.2 (1Es) 27.4 (1F2) 53.4 (2F2)	-0.6
	13.685	21/10/91 09:11	<u>28</u> 44 57	2	29 (1F) 52 (2F)	1	13.9 (1Es) 67.8 (3F2) 28.0 (1F2) 54.0 (2F2)	0
	13.685	17/01/92 09:30	<u>28</u>	2	28 (1F) 50 (2F)	0	27.8 (1F2)	-0.2
ARGA	11.79	09/10/91 08:40	<u>26</u> 47	3	26 (1F) 49 (2F)	0	28.1 (1F2)	2.1
	11.79	17/01/92 09:25	<u>24</u> 45	2	25 (1F) 46 (2F)	1	23.2 (1F2) 46.3 (2F2)	-0.8
WERT	17.845	21/10/91 13:48	<u>26</u>	2	29 (1F)	3	26.7 (1F2)	0.7
	17.845	17/01/92 09:40	<u>21</u>	2	24 (1F)	3	23.7 (1F2)	2.7

Tableau 3: Comparaison des angles de site mesurés avec les directions d'arrivée calculées, à l'aide du modèle de propagation ASAPS (méthode 1) et des transpositions en oblique des ionogrammes verticaux de la station ionosphérique de Lannion (méthode 2). Les fréquences des émetteurs sont comprises entre 12 et 22 MHz.

TX	FREQUENCE (en MHz)	DATE HEURE (T.U.)	SITES LOSQUET en deg.	PAS en deg.	SITES METHODE 1 en deg.	ECART en deg.	SITES METHODE 2 en deg.	ECART en deg.
WIEN	6.155	08/10/91 09:43	<u>10</u> 50 77	3	14 (2E)	4	4.4 (1E) 29.1 (3Eh) 14.9 (2E) 44.5 (3F2) 23.9 (3Eb)	4.9
LENK	9.535	22/10/91 13:54	24 <u>53</u>	3	13 (1E) 33 (1F) 52 (2F)	-1	12.1 (1Eb) 49.8 (2F2) 16.0 (1Eh) 25.6 (1F1)	-3.2
ALLO	6.17	08/10/91 14:48	<u>22</u> 45 75	5	28 (1E) 52 (1F)	6	25.7 (1E) 75.3 (3F2) 44.4 (1F1) 67.9 (2F2)	3.7
SKEL	9.76	22/10/91 14:50	<u>33</u>	3	40 (1F) 60 (2F)	7	28.4 (1F2) 53.8 (2F2) 65.5 (3F2)	-4.6
	9.76	17/01/92 09:25	<u>34</u> <u>59</u>	2	37 (1F) 58 (2F)	-1	28.2 (1F2) 50.6 (2F2) 62.6 (3F2)	-3.6
JUNG	6.09	08/10/91 09:22	<u>12</u> 51	5	15 (1E) 28 (1F) 57 (2F)	3	11.0 (1E) 47.2 (1F2) 59.9 (2F2)	-1

Tableau 4: Comparaison des angles de site mesurés avec les directions d'arrivée calculées, à l'aide du modèle de propagation ASAPS (méthode 1) et des transposition en oblique des ionogrammes verticaux de la station ionosphérique de Lannion (méthode 2). Les fréquences des émetteurs sont comprises entre 6 et 12 MHz.

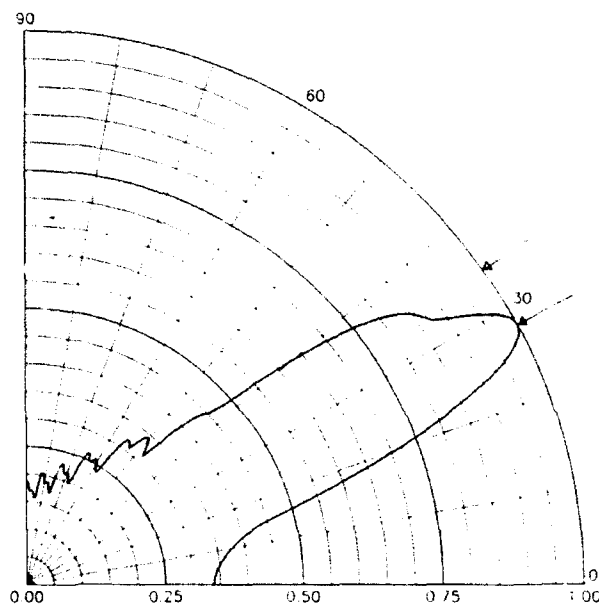


Figure 7 : Diagramme de réception en site simulé  
Fréquence = 15.45 MHz  
Azimut de visée = 86°  
Trajet n°1 : Elévation introduite = 28°  
Trajet n°2 : Elévation introduite = 35°

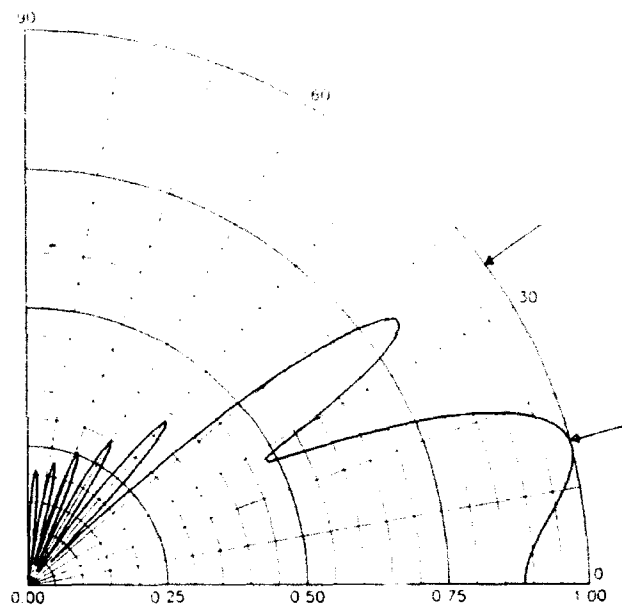


Figure 8 : Diagramme de réception en site simulé  
Fréquence = 15.45 MHz  
Azimut de visée = 86°  
Trajet n°1 : Elévation introduite = 15°  
Amplitude = 1  
Trajet n°2 : Elévation introduite = 35°  
Amplitude = 1

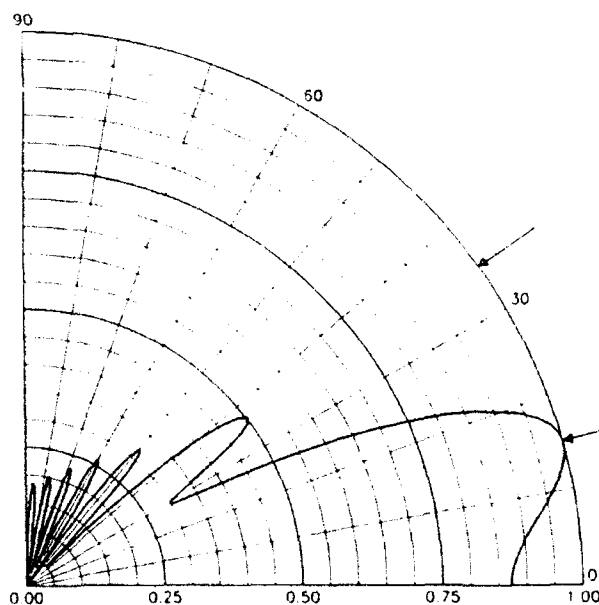


Figure 9 : Diagramme de réception en site simulé  
Fréquence = 15.45 MHz  
Azimut de visée = 86°  
Trajet n°1 : Elévation introduite = 15°  
Amplitude = 1  
Trajet n°2 : Elévation introduite = 35°  
Amplitude = 0.5

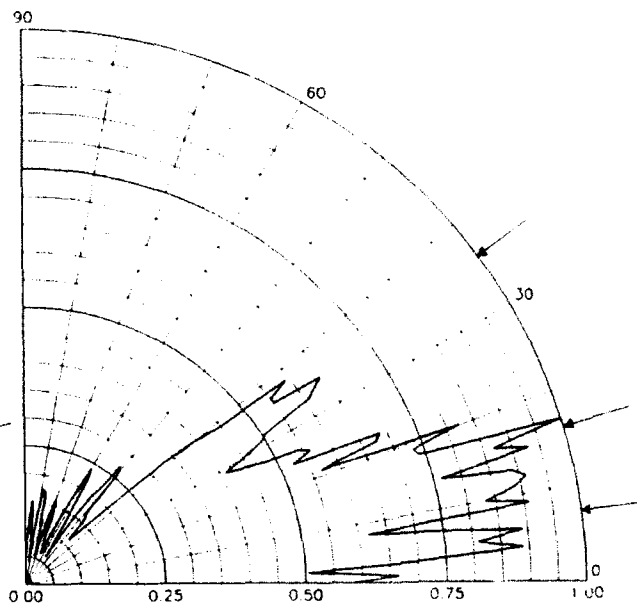


Figure 10 : Diagramme de réception en site simulé  
Fréquence = 15.45 MHz  
Azimut de visée = 86°  
Trajet n°1 : Elévation introduite = 7°  
Amplitude = 0.7  
Trajet n°2 : Elévation introduite = 15°  
Amplitude = 1  
Trajet n°3 : Elévation introduite = 36°  
Amplitude = 1

## DISCUSSION

**E. SCHWEICHER**

Avez-vous employé une apodisation pour réduire les lobes secondaires de votre réseau d'antennes?  
*Have you used apodization to reduce the side lobes of your array?*

**AUTHOR'S REPLY**

Une fois le lobe formé, à l'aide de la technique que nous avons présentée, nous ne mettons pas en oeuvre d'autres méthodes.  
*Once the lobe has been formed, using the technique we have described, we do not use any other methods.*

**J. BELROSE**

When you compared measured with predicted elevation angle of arrival of short wave broadcast signals, have you included in your analysis the vertical plane patterns of the transmitting antennas?

**AUTHOR'S REPLY**

Nous ne tenons compte que du diagramme des antennes du réseau de réception et nous considérons qu'un mode de propagation est dominant lorsqu'il est mesuré comme tel.

*We only take into account the receiver array antenna pattern and we consider that a propagation mode is dominant when it is measured as such.*

**D. CHOI**

I recall your measurement showing ~ 5° beamwidth at 20 MHz. Fundamentally, an array is limited in beamwidth by  $\lambda/D$  in radians where  $D$  = size of the aperture and  $\lambda$  is the wavelength. At 20 MHz,  $\lambda/D = 15/68$  and the best beamwidth is  $\approx 12^\circ$ . Can you comment on this?

**AUTHOR'S REPLY**

The radius of the antenna array is 68.8 m but the diameter is about 140 m which gives the shown beamwidth.

**E. SCHWEICHER**

Sur votre second transparent, j'ai noté la présence (à partir du calculateur) de 2 bus: un bus GPIB (car je croyais que GPIO était une faute de frappe) et un bus IEEE que je croyais identiques. Quelle est la différence entre ces deux bus? Si GPIO est réellement GPIO et non GPIB, que signifie GPIO?

*On your second viewgraph, I noted that there were 2 buses (starting from the computer), a GPIB bus (as I think GPIO was a typing error) and a IEEE bus, which I thought were identical. What is the difference between these 2 buses? If GPIO is really GPIO and not GPIB what does GPIO stand for?*

**AUTHOR'S REPLY**

Le bus GPIO comme le bus IEEE est un bus parallèle mais contrairement au bus IEEE il n'y a aucun échange de protocole d'où un gain de rapidité au niveau du transfert des données.

*Like the IEEE bus, the GPIO bus is a parallel bus, but unlike the IEEE bus there is no exchange of protocol, which makes for more rapid data transfer.*

**C. GOUTELARD**

Commentaire:

Je voudrais faire remarquer, suite à la question du Dr Belrose, qu'il n'y a pas à se préoccuper du diagramme de rayonnement de l'émetteur. En effet, celui-ci produit des trajets divers en fonction de son diagramme. Si, par exemple, il en supprime un, ce n'est pas important: le récepteur ne traitera que les trajets par lesquels les ondes se seront propagées, et le traitement sera plus simple. Le rôle du récepteur ici est de déterminer les trajets qui se sont propagés et non ceux qui auraient pu se propager.

*Following Dr. Belrose's question, I would like to point out that we don't need to concern ourselves with the emitter's radiation pattern. In effect, the emitter produces different paths depending on its pattern. If it cancels one, for example, this is not important; the receiver will only process those paths along which waves are propagated, and the processing will be simpler. In this case, the role of the receiver is to determine the paths which have been propagated and not those which might have been propagated.*

## TEST RESULTS OF THE ADVANCED TRANSLATOR PROCESSING SYSTEM

Alison K. Brown  
 NAVSYS Corporation  
 14960 Woodcarver Road  
 Colorado Springs, CO 80921  
 USA

William Sward, NAVSYS Corporation  
 Peter Brown, NAVSYS Corporation

### 1. SUMMARY

The Advanced Translator Processing System was developed to provide a low cost alternative to tracking translated GPS signals. It is backwards compatible with the Translator Processing System (TPS) currently used to track Ballistic Missile Translators (BMT) by the US Ranges. NAVSYS has also developed a BMT compatible Translator.

A key element of the ATPS is the Preamplifier/Downconverter (P/DC) module which was developed by NAVSYS to condition the received Translator signal so it can be tracked with a conventional C/A code receiver. This significantly reduces the cost of the hardware by allowing the ATPS to be constructed mostly from off-the-shelf components. The ATPS and Translator were tested against the RAJPO Ballistic Missile and Translator Processing System Specifications. This paper presents the results of tests demonstrating the performance of the ATPS and Translator under different operating conditions.

NAVSYS has also developed a Post Test Processing System (PTPS), incorporating the same components, which records the Translator data during the tests. When this data is played back, high-accuracy Time and Space Position Information (TSPI) data can be computed on the vehicle even under extreme dynamics.

### 2. INTRODUCTION

#### 2.1 Current Translator Systems

GPS Translator tracking systems have been developed by the US Navy and by the tri-service Range Applications Joint Program Office (RAJPO). These are designed to provide real-time Time Space Position Information (TSPI) data for high dynamic flight test and tracking applications.

The GPS Translator concept is illustrated in Figure 1. The Translator installed on-board the missile converts the received L-band GPS signals to S-band and rebroadcasts them to a ground-based Translator Processing System (TPS) where the location of the vehicle is determined. The RAJPO BMT Translator signal spectrum is illustrated in Figure 2. To limit the S-band signal bandwidth, only the C/A code GPS signals are translated and broadcast to the TPS.

#### 2.2 NAVSYS Advanced Translator Processing System

For many of the potential applications for GPS Translators, the size and cost of the Translator Processing System currently in use in the United States is prohibitive. NAVSYS has developed a smaller, low cost Advanced Translator Processing System (ATPS) suitable for use in these applications.

The reduction in size and cost of the system was achieved through the use of off-the-shelf components where possible. The key element of the ATPS is the Preamplifier/Downconverter (P/DC) subsystem. The P/DC tracks the Pilot Carrier from the GPS Translator and uses this signal to control downconversion of the received S-band signal back to the L-band signal received by the Translator. This signal is then passed to a conventional GPS receiver for processing.

The performance of the ATPS is limited only by the tracking performance of the GPS receiver. Testing has been performed using a number of GPS receivers including the Mini-Rogue SNR 8A, the Ashtech XII, and the Navstar XR3-P receiver.

### 3. NAVSYS GPS TRANSLATOR

The NAVSYS GPS Translator was designed to be backwards compatible with the Ballistic Missile Translator (BMT) Type A, with the following additional capabilities:

While the BMT is designed to operate at one of 89 possible frequencies, the frequency of the NAVSYS Translator is field selectable to any of the 89 channels through an internal switch. The NAVSYS Translator is also designed to allow BPSK data to be modulated on the Pilot Carrier (PC) signal. This enables the Translator to be used as a telemetry link as well as a tracking system. The Translator will also allow the S-band GPS signal to be modulated with a pseudo-noise M-code. This feature can be used to allow the CDMA operation of multiple translators sharing the same frequency band.

The Translator specification is summarized in Table 1.

### 4. PREAMPLIFIER/DOWNCONVERTOR

The P/DC is designed to operate with signals from either the BMT or the NAVSYS Translator at any of the 89 S-band frequency channels. The P/DC converts the received Translator signals from S-band for subsequent processing by a GPS receiver. The P/DC specifications are summarized in Table 2.

The P/DC follows an S-band preamplifier, normally a range preamplifier providing  $25 \pm 5$  dB of gain with 2 dB noise figure. The S-band signal from the preamplifier is mixed to an IF frequency and AGC amplified. The leveled signal is divided, one path being bandlimited to the C/A-code spectrum, translated to L-band and output to the GPS receiver, the other being translated to baseband and output to the P/DC Controller.

The P/DC Controller provides control voltages to set the AGC gain and the synthesizer reference frequency

Table 1 Translator Specification

<b>Input Signal:</b>	GPS L1 C/A Code
Frequency	1575.42 MHz $\pm$ 50 kHz
Signal level	-145 dBm to -110 dBm
<b>Output Signal:</b>	Channel Field Selectable N = 1 to 89
Center Frequency	2201.92 + N MHz
PC Frequency	2200 + N MHz
RF Power Level	+35 $\pm$ 2 dBm (3 watts)
Test Option	+30 $\pm$ 2 dBm (1 watt)
PC Level	RF Power Level - (10 $\pm$ 2 dB)

<b>Modulation Options:</b>	
M-Code (CDMA)	DC to 1 MBPS
PC-Data (Telemetry)	DC to 62.5 KBPS

**Performance Requirements:**Phase Stability

Condition	Requirement
10 second period	$\leq 30^\circ$
10 dB input change	$\leq 5^\circ$
-145 to -110 dBm input range	$\leq 20^\circ$

Frequency Accuracy  $\pm 10$  PPMFrequency Drift 0.1 sec Allan Variance  $< 10^{-10}$ Pilot Carrier Phase Jitter

Double Sided Noise Bandwidth	RMS Phase Jitter
1 to 10 Hz	$\leq 8^\circ$
1 to 1500 Hz	$\leq 10^\circ$
1 to 15,000 Hz	$\leq 15^\circ$

Output Phase Jitter

Double Sided Noise Bandwidth	RMS Phase Jitter
1 to 10 Hz	$\leq 7^\circ$
1 to 1500 Hz	$\leq 15^\circ$

**Physical Requirements:**

<u>Immersed Volume</u>	25 cubic inches
<u>Weight:</u>	1.8 pounds (excluding antennas, cabling, connectors, and primary power source)
<u>Power</u>	20 to 32 VDC, 22 watts

(provided from a voltage controlled crystal oscillator (VCXO)).

The P/DC Controller is a DSP computer card embedded in a host IBM Personal Computer. The P/DC Controller digitally samples the Pilot Carrier signal at 200 kHz and processes this data to provide the AGC and VCXO control signals. The AGC control is required to adjust the received signal level of the translated GPS signal to the nominal level expected by the GPS receiver. The VCXO control is required to remove the frequency offset on the S-band carrier from the Translator clock. This function realigns the L-band code and carrier Doppler shifts so that the receiver can function without any modification. The L-band signal then appears to the GPS receiver as though it were received along the path translator-satellite-ground.

**5. TEST RESULTS**

The Translator and ATPS system test results are described in the following sections. Component level tests were performed to verify that the Translator and P/DC met their individual specifications, the system tests were then used to verify the performance of the systems under laboratory conditions.

**5.1 Navigation Performance Tests**

The test configuration shown in Figure 3 was used to verify the operation of the Translator and P/DC under static conditions and to compare the navigation performance to that with a reference GPS receiver.

Table 2 P/DC Specification

**S-Band Signal Input**

Composite of Translator output signals	
Center Frequency:	2201.92 MHz to 2390.92 MHz (selectable in 1 MHz steps)
PC Frequency:	Center Frequency + PC Offset
RF Level:	-97 dBm to -21 dBm
PC Level:	10 $\pm$ 2 dB below RF level
Preamp Noise Level:	-147 $\pm$ 5 dBm/Hz
PC Offset:	-1.92 MHz + M MHz M = 1 to 56

**Signal Encoding**

M-Code	DC to 1 MBPS
--------	--------------

**L-Band Signal Output**

Translator L-Band input, Translator input noise, and amplifier thermal noise	
Center Frequency:	1575.42 MHz
RF Level:	-94.5 dBm to -77.5 dBm

**Performance Requirements**

<u>Attenuator Control</u>	Power: 0 to 70 dB Distance 1:2000
<u>Frequency Control</u>	$\pm 10$ ppm $\pm 20$ kHz at S-Band
<u>Frequency Accuracy</u>	0.01 PPM after 3 minute warm-up
<u>Frequency Drift</u>	0.1 sec Allan Variance $< 10^{-10}$
<u>Pilot Carrier Channel Phase Jitter</u>	
Double Sided Noise Bandwidth	RMS Phase Jitter
1 to 10 Hz	4 $\circ$
1 to 1500 Hz	5 $\circ$
1 Hz to 15 KHz	8 $\circ$
<u>Signal Channel Phase Jitter</u>	
Double Sided Noise Bandwidth	RMS Phase Jitter
1 to 10 Hz	4 $\circ$
1 to 1500 Hz	8 $\circ$

**Physical Requirements**

<u>Immersed Volume</u>	(46 cubic inches) All dimensions exclusive of connectors and mountings.
<u>Weight</u>	3.3 pounds MAX, excluding antennas, cabling, connectors, and primary power source.
<u>Power</u>	20 to 32 VDC (23 watts)

Initial measurements were made between two XR3-P receivers directly tracking the GPS satellite signals (Figure 4). The double difference delta-ranges (DDR) between two satellites and two receivers operating from a common antenna was used to characterize the phase-noise on the tracking loops. The variance of this phase-noise is equivalent to twice the individual phase-noise variance on a single satellite.

A second set of results were made between a reference receiver and a receiver tracking the signals from the GPS Translator (Figure 5). The results showed a slight degradation in signal to noise ratio in the translator signal. There was no noticeable increase in the carrier phase noise observed from the DDR measurements. In both cases the DDR 1-sigma was around 5 mm, indicating that the receiver tracking loop noise of roughly 2.5 mm is dominating any phase noise introduced by the Translator and P/DC.

The relative position and velocity accuracy was also very similar between the Translator and dual-receiver test



cases. The velocity accuracy was 0.01-0.02 m/s 1-sigma in both cases for PDOPs of 2.5 or better.

### 5.2 Power Level Tests

The purpose of the power test was to demonstrate the ability of the AGC loop to track the translated signals throughout the full range of operation of the P/DC. To perform this test, variable attenuation was introduced into the path from the Translator to the P/DC to vary the received S-band RF power from -116 dBm to -56 dBm, equivalent to tracking a Translator over a 1:1000 change in range.

Throughout the tests, the SNR through the Translator was degraded by 0 to 3 dB. The DDR measurements provided a 1-sigma variance of around 5 mm, except where steps were observed due to a timing error problem in the XR3-P data. The observed velocity in the high dynamic mode was between 1-2 cm/s 1-sigma. The relative position error between the two GPS receivers was generally within 3 meters.

### 5.3 Simulated Dynamics Tests

Since the ATPS testing was performed in the laboratory, it was obviously not possible to demonstrate the performance under actual vehicle dynamics. Instead, simulated Doppler shifts (or oscillator frequency shifts) were introduced on the Translator signal output by varying the frequency of the Translator oscillator. By heating and cooling the Translator oscillator, it was possible to cause the frequency to vary by over 10 ppm. This is equivalent to a frequency shift at S-band of 20 kHz or a Doppler velocity shift of 3000 m/s. This test verified the ability of the P/DC Controller to remove this frequency shift without disrupting the operation of the ATPS receiver.

The change in frequency inserted an apparent acceleration into the translated GPS signal of up to 10 m/s/s (1 g) (Figure 6). If this were not corrected for in the P/DC, the frequency offset would cause an error in the DDR measurements and the velocity computed by the XR3-P receiver. In addition, the XR3-P receiver would have difficulty maintaining code lock due to the difference in the apparent code and carrier Doppler shifts. The P/DC correction removed the Translator clock offset frequency so the L-band signal into the XR3-P receiver was unaffected by the apparent Doppler shift.

During this test, the Translator SNR was within 2 dB of the reference receiver, the variance of the DDR measurements was less than 6 mm, and the velocity provided by the receiver was accurate to 4 cm/s. The velocity accuracy was the only noticeable change in this test. Under "static" conditions the velocity accuracy was better than 2 cm/s, while under the "dynamic" conditions the accuracy degraded to 4 cm/s, with the receiver in high-dynamic mode. However, this performance is excellent when tracking a signal that varies by 10 m/s/s, and it demonstrates the precision of the P/DC tracking loops in removing the S-band Doppler frequency offset.

### 5.4 Interference Tests

These tests demonstrated the ability of the P/DC to track individual Translator signals when more than one Translator is operating, and to switch between the Translators. In the first test, two Translators were connected to the input of the P/DC with their frequencies offset by 4 MHz (2206.5 MHz and 2210.5 MHz). The two Translators were set with a relative power level of 6 dB. The P/DC was able to track both signals accurately throughout the test.

A second test was performed with the Translators separated in frequency by 20 MHz and the relative power level adjusted to determine the minimum and maximum tracking power level for the P/DC. The P/DC was able to track both Translators and provide DDR and velocity accuracies comparable to the tests described in 5.2 (3-4 mm and 2 cm/s) over a 49 dB operating range. When the signal power for the low power Translator was dropped further, the P/DC receiver had trouble maintaining lock on the signal.

## 6. FURTHER DEVELOPMENTS

NAVSYS is currently developing a Post-test Processing System (PTPS) based on the ATPS design.

The PTPS is illustrated in Figure 7 and consists of a single rack of equipment and a host SUN computer that is designed to record and post-test process Translator data.

In record mode, the PTPS records data from two tracking antennae (typically one Left Hand Circularly Polarized (LHCP) the other RHCP Translator signals from the same vehicle) and one static reference GPS antenna to a high speed Very Large Data Store (VLDS) recorder.

In post-test mode, the PTPS permits playback of the recorded data and recovery of the GPS data. The specification of the Post Test Processing System is listed in Table 3.

### 6.1 Data Recording

The configuration of the PTPS recording system is shown in Figure 7.

The PTPS system uses two P/DC units and controllers, as described for the ATPS system, to AGC the incoming Translator signals and convert the signal back to L-band. The resulting L-band GPS signals are fed, along with a reference channel signal from a static antenna/preamplifier, to three Predetection GPS Receivers (PGR). The PGR is a NAVSYS designed GPS receiver that outputs In-Phase and Quadrature samples of the incoming signal at 2 MSps.

The resulting data streams are packed into an 8-bit word along with digitized Pilot Carrier data from the two P/DCs and passed to a Metrum VLDS Data recorder recording at 2 MBytes/sec.

Timing of the entire system comes from a highly stable ovenised 10 MHz oscillator.

The two P/DC controller DSP cards are embedded in a SUN Sparcstation2 host computer that also controls the tape recorder.

### 6.2 Post-Test Processing

The configuration of the PTPS post-test system is shown in Figure 8. The recorded data is retrieved from tape into an interface card. The interface card buffers 1 second of data and routes the desired GPS channel to two arrays of up to 10 accelerator chips. Each accelerator chip is instructed by the host to track one of the visible satellites. The accelerator chip performs code correlation and tracking under control of one of two DSP cards. Aiding data, if available, is passed from the host to the DSP to control the tracking algorithms. Acquisition over a wide range of Dopplers can be achieved using search bins of 500 Hz and repeatedly processing the data in the 1-second data buffer. Once signal lock has been acquired, the tape will start streaming and processing can be achieved in real-time. Loop logic permits real-time aiding between

Table 3 PTPS System Specification

<b>Input</b>	LHCP and RHCP Translator signals at S-band -76 dBm to -16 dBm L1 reference signal per ICD-GPS-200
<b>Output to tape</b>	Composite words of 1-bit digitized (I&Q) 2 MSps GPS signals (3 ea) and 5-bit digitized 400 kHz Pilot Carrier signals (2 ea)
<b>Tape Recorder</b>	Metrum VLDS
Transfer rate	2 MByte/sec sustainer transfer rate
Tape medium	T-120 video cassette
Record time	84 minutes
<b>Host computer</b>	SUN Sparestation 2 with two C-30 DSP cards.
<b>Physical characteristics (without host computer)</b>	
Weight	160 lbs
Dimensions	26" x 17" x 24"
<b>Maximum Dynamics for PTPS Tracking Loops:</b>	
Velocity	6000 m/s
Acceleration	50 g
Jerk	50 g/s
Maximum Doppler	$\pm 92$ kHz
<b>PTPS Tracking Loop Performance</b>	
<u>Signal L<sub>1</sub> C/A code</u>	
Performance	40 dB-Hz (SNR) PR @ 1 Hz = 6 m (1-sigma) DR @ 1 Hz = 1 cm (1-sigma)
<u>Acquisition Threshold</u>	
<u>Tracking Threshold</u>	
AFC Tracking	21 dB-Hz @ 10 g @ 10 g/s
Phase Tracking (data aided)	22 dB-Hz @ 5 g @ 5 g/s

channels to assist reacquisition if lock is lost. In this way, the LHCP can aid the RHCP channel (and vice versa) during signal fades. Host control permits tape rewind and reacquisition if lock is lost on both channels during streaming.

A preliminary processing pass using the static reference receiver data provides navigation data, which can also be used to navigation data-aid the tracking loops. This results in roughly a 6 dB improvement in the tracking loop threshold.

## 7. CONCLUSIONS

The results of the ATPS system tests are summarized in Table 4. The development of both the Translator and the P/DC proved to be highly successful and all of the design objectives were met.

The Translator demonstrated its backward compatibility with the BMT, and its capability of field-selectable operation, at all 89 of the BMT frequencies.

The Advanced Translator Processing System passed all of the system tests and demonstrated that the P/DC is capable of tracking GPS signals with no noticeable performance degradation over a conventional GPS receiver. Carrier phase measurements were provided to an accuracy of 3 mm and velocity was determined to 1 cm/s (1-sigma) with the receiver in high-dynamics mode of operation. The P/DC was capable of tracking translated

Table 4 Summary of ATPS System Test Results

<b>Navigation Performance</b>	
SNR Degradation	1 dB
DDR-Observed Phase Noise	4 mm (1-sigma)
Position Error	< 3 meters
Velocity Error	2 cm/s (1-sigma)
<b>P/DC Power Level Thresholds</b>	
Maximum	-56 dBm
Minimum	-116 dBm
<b>Frequency Tests</b>	
Full Range	$\pm 10$ ppm ( $\pm 3000$ m/s)
<b>Simulated Dynamics</b>	
	0 to 10 ppm (10 m/s/s acceleration)
<b>Interference Tests</b>	
Translators separated by 4 MHz and 6 dB - Passed	
Translators separated by 20 MHz and 49 dB - Passed	

GPS signals to this precision over a 60 dB range of power, providing a range of operation of 1:1000 in distance.

The P/DC has also been integrated into a Post-Test Processing system that is capable of recording the Translator signals during a mission. The PTPS is designed to replay this data and make high accuracy measurements throughout the vehicle's trajectory even under high dynamics.

## 8. REFERENCES

1. J.B. McConnell, R.H.Greenberg, R.B.Pickett, P.C. Wildhagen, A.K.Brown, "Advances in GPS Translator Technology," ION Satellite Division Conference Proceedings, September 1989.
2. "Miniaturized GPS Translator and Translator Processing System Development, Phase II Final Report," NAVSYS Document TPS-91-05, August 1991.

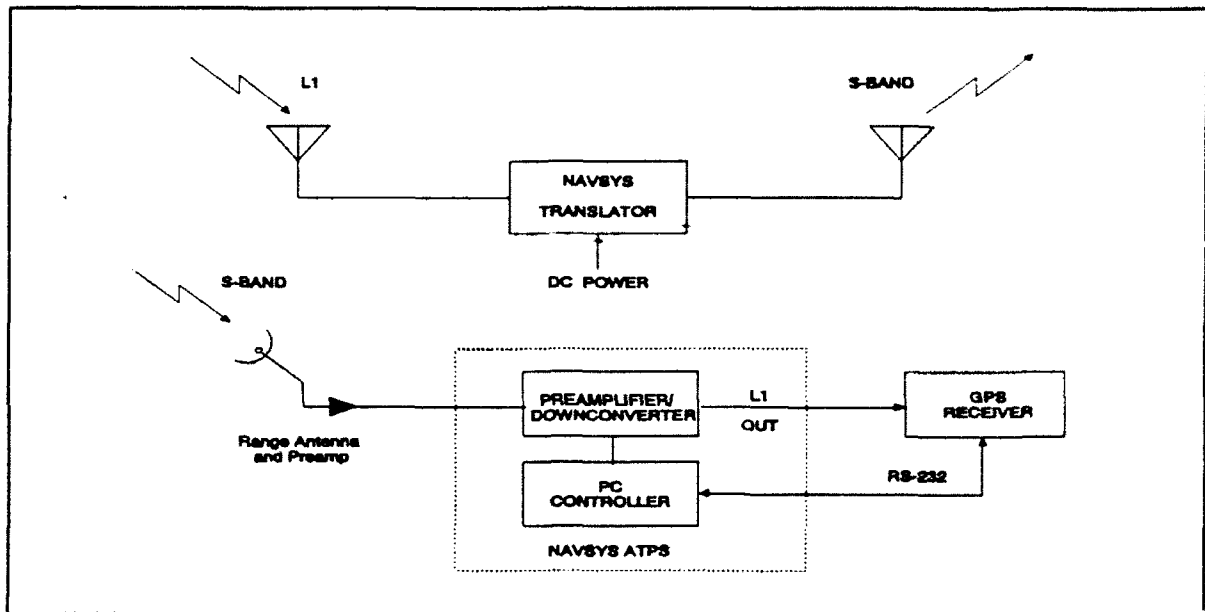


Figure 1 GPS Translator and Advanced Translator Processing System

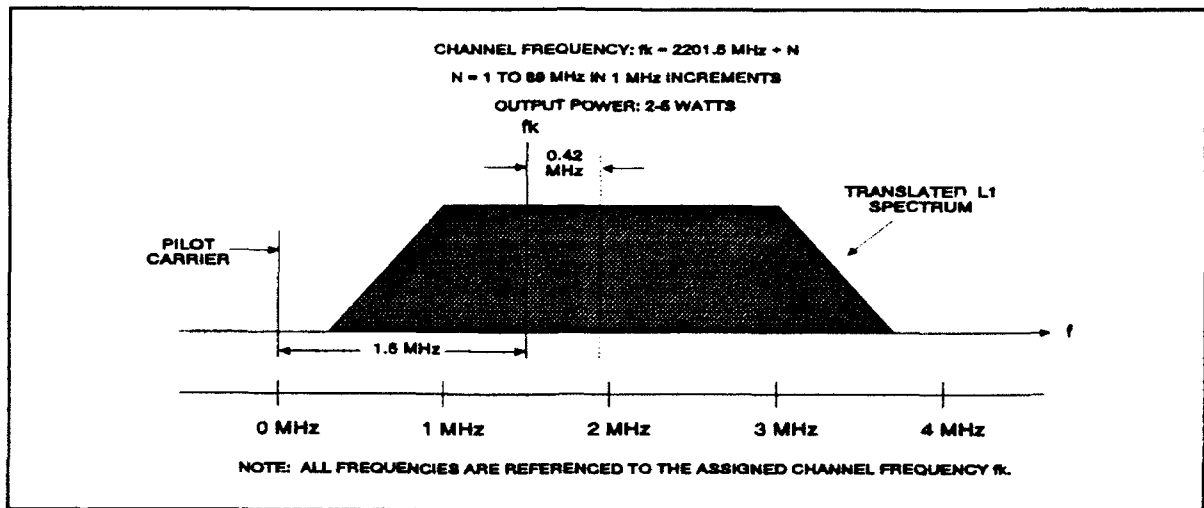


Figure 2 BMT Translator Signal Spectrum

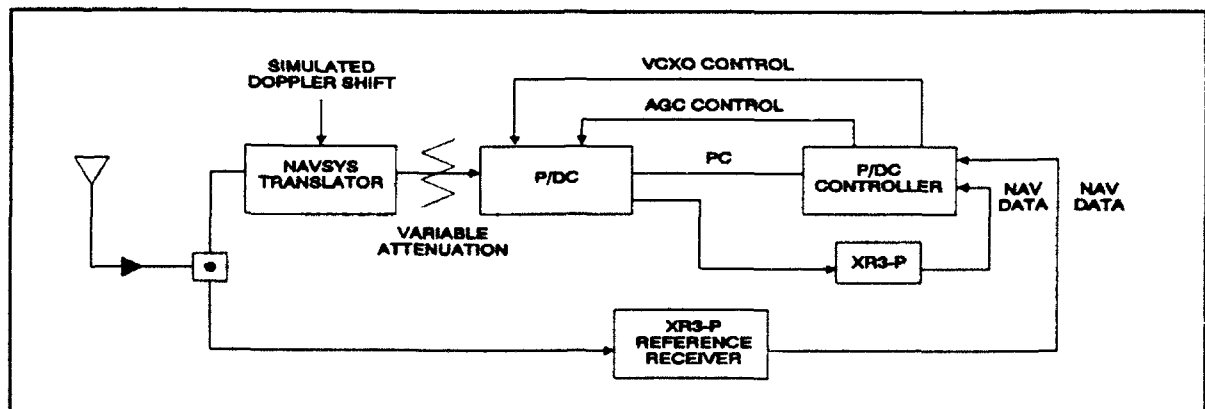


Figure 3 System Test Set-Up

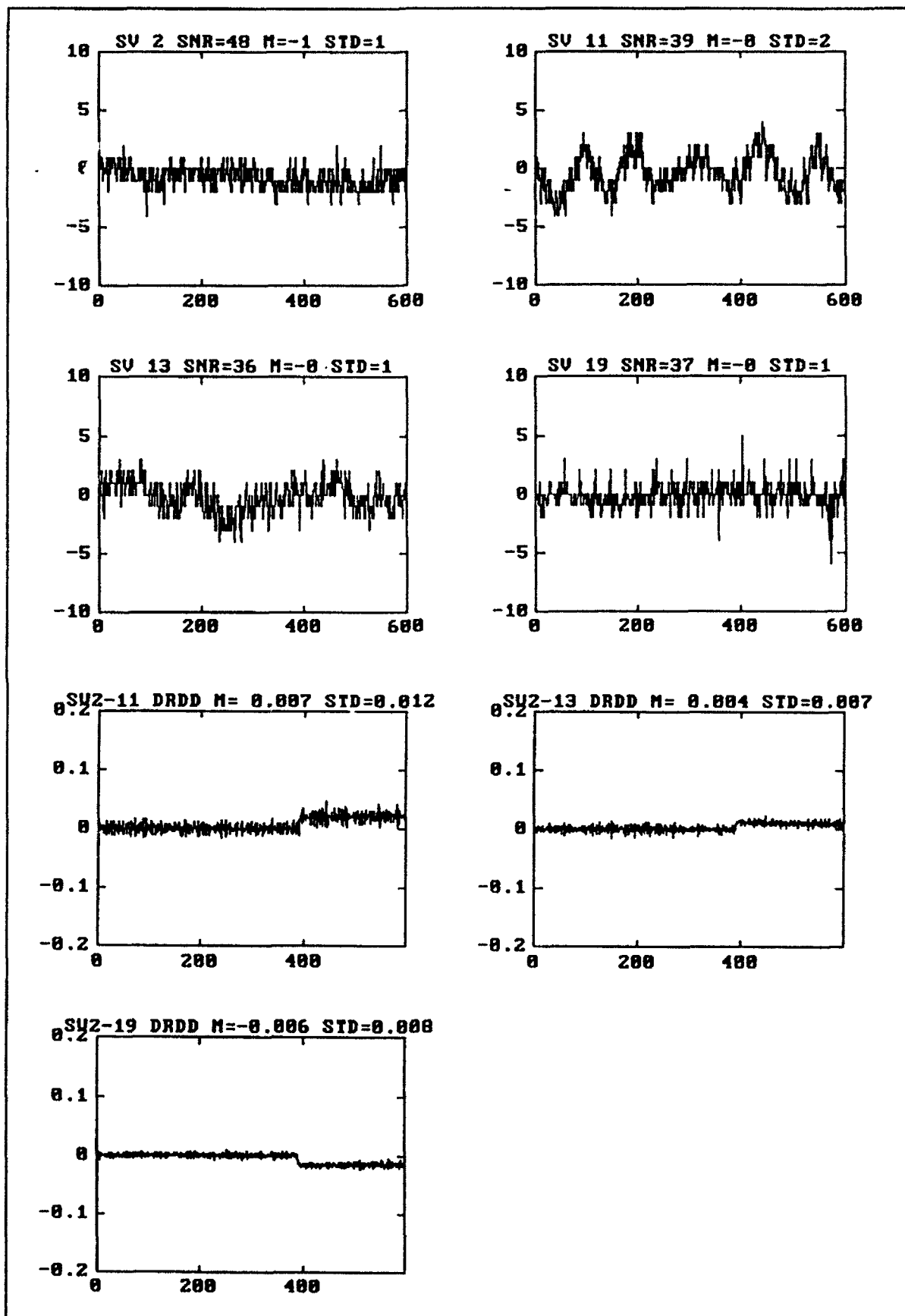


Figure 4 XR3-P vs XR3-P Comparison

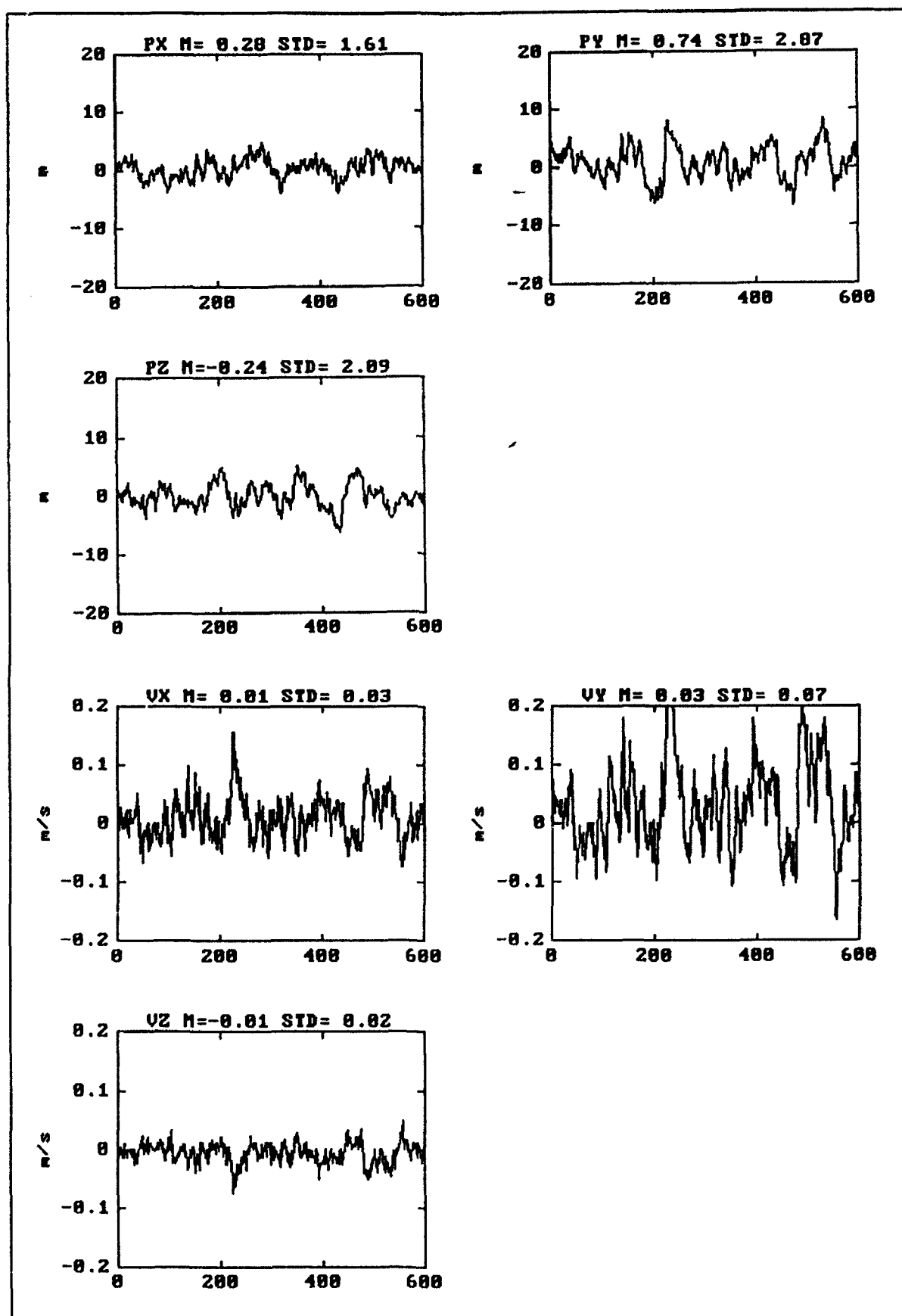


Figure 4 XR3-P vs XR3-P Comparison (Continued)

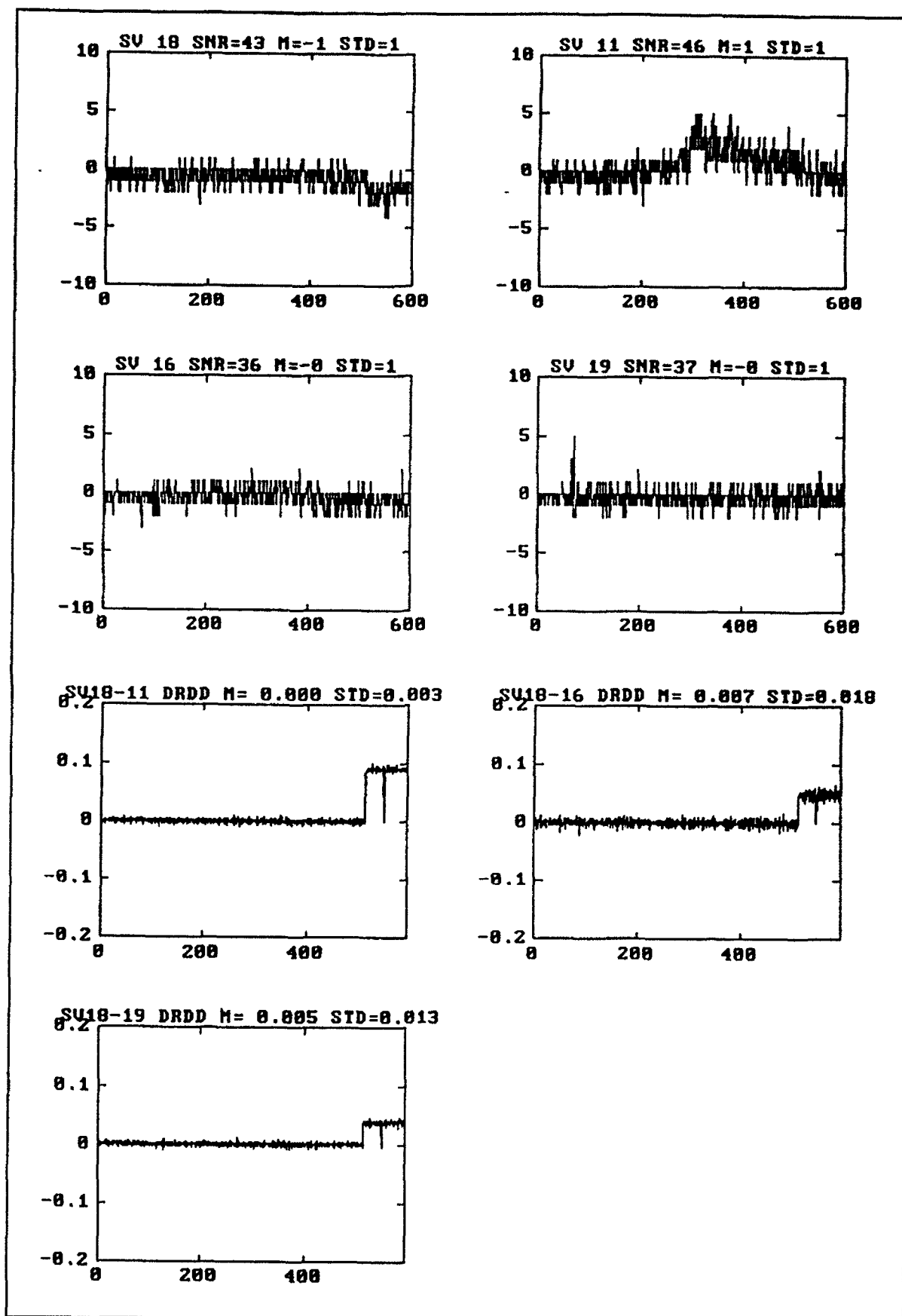


Figure 5 Navigation Performance Tests

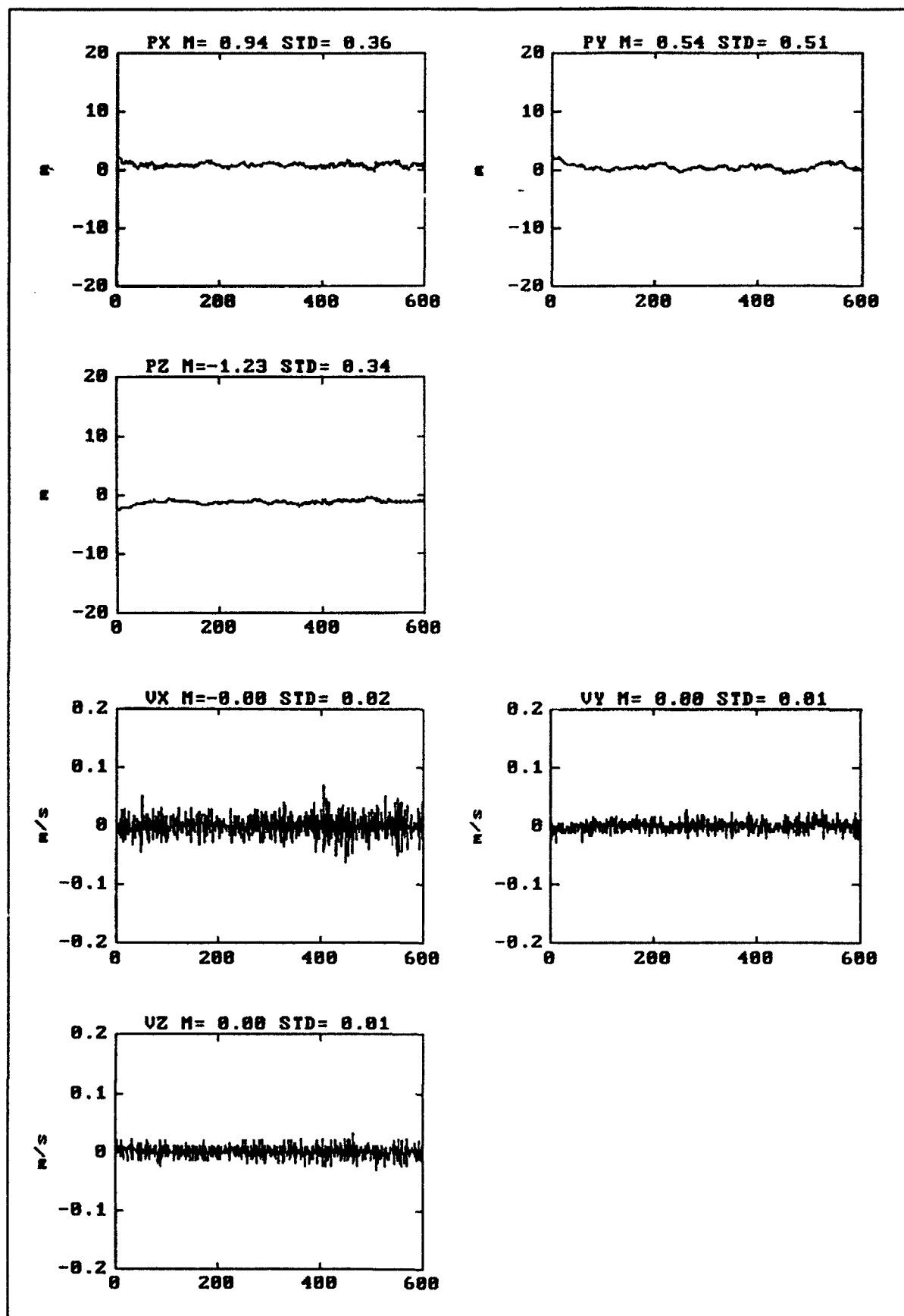


Figure 5 Navigation Performance Tests (Continued)

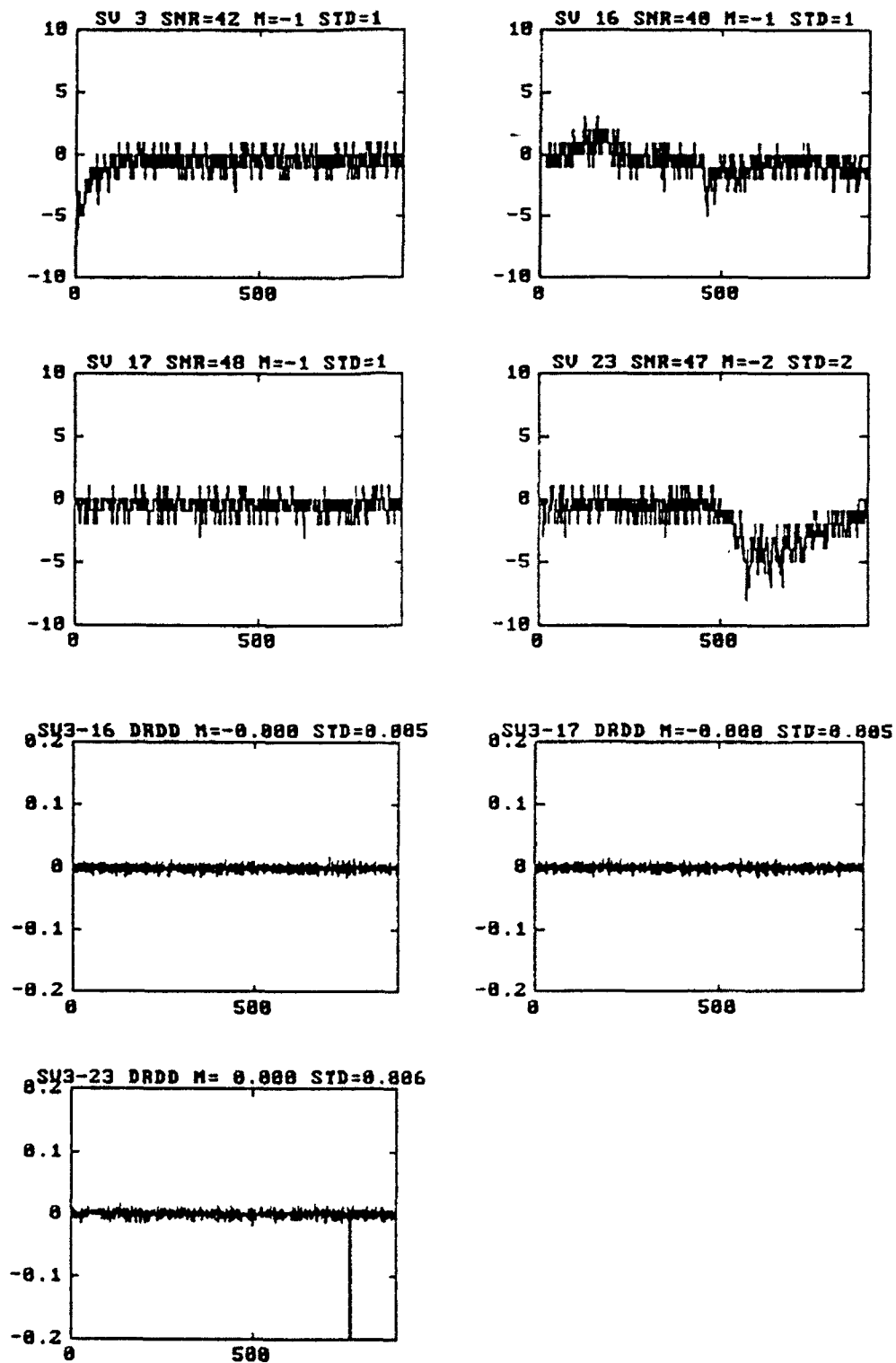


Figure 6 Simulated Dynamics Tests



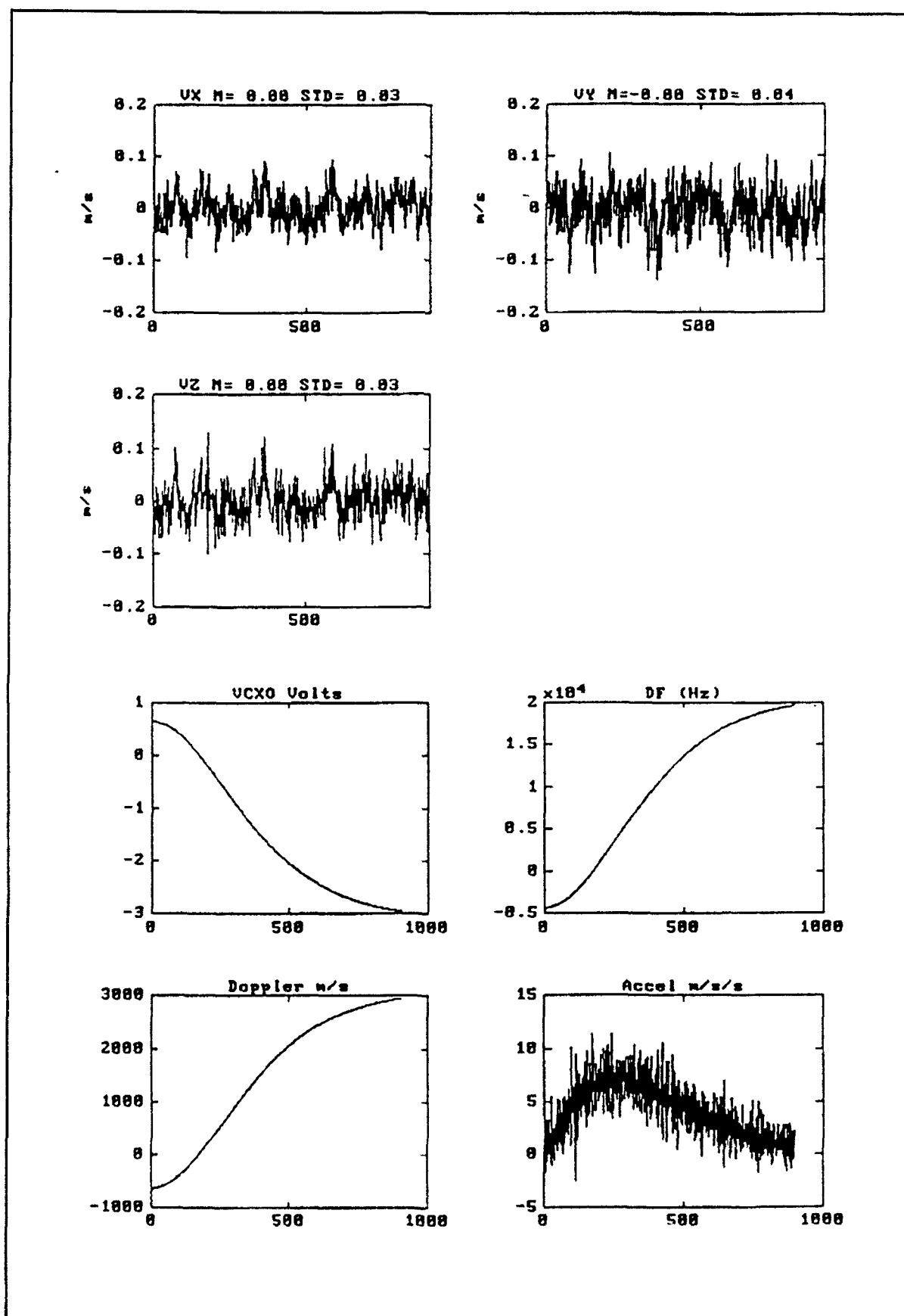


Figure 6 Simulated Dynamics Tests (Continued)

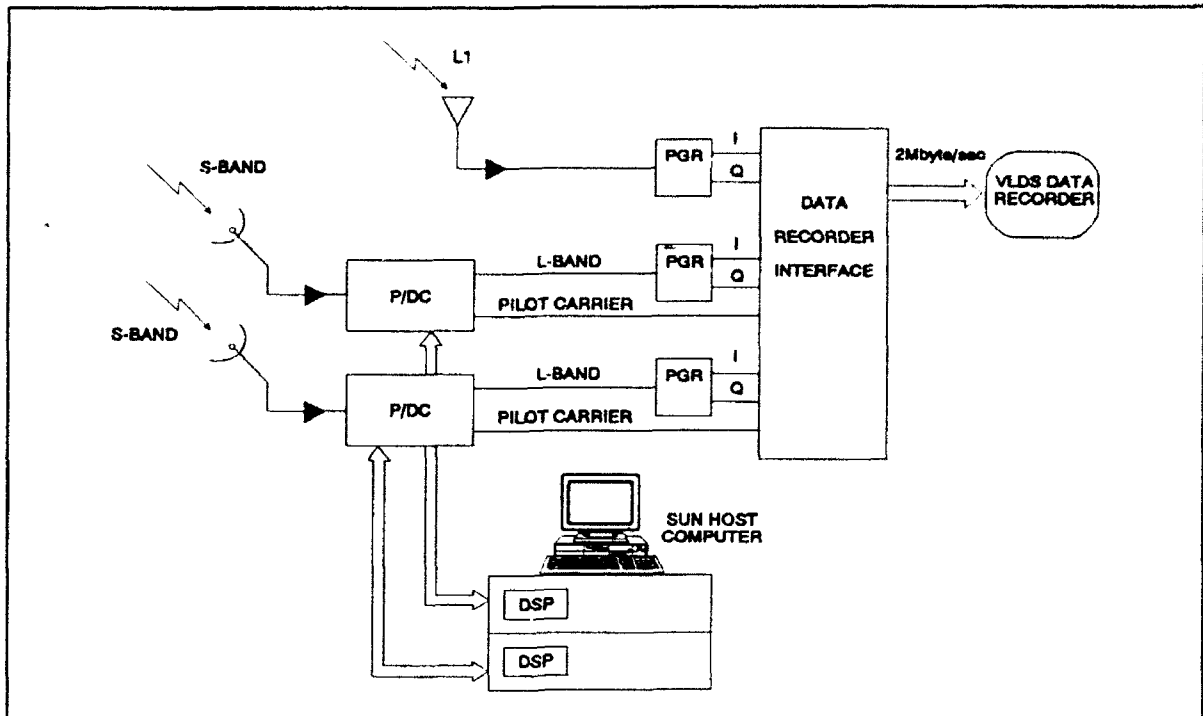


Figure 7 PTPS Recording Mode

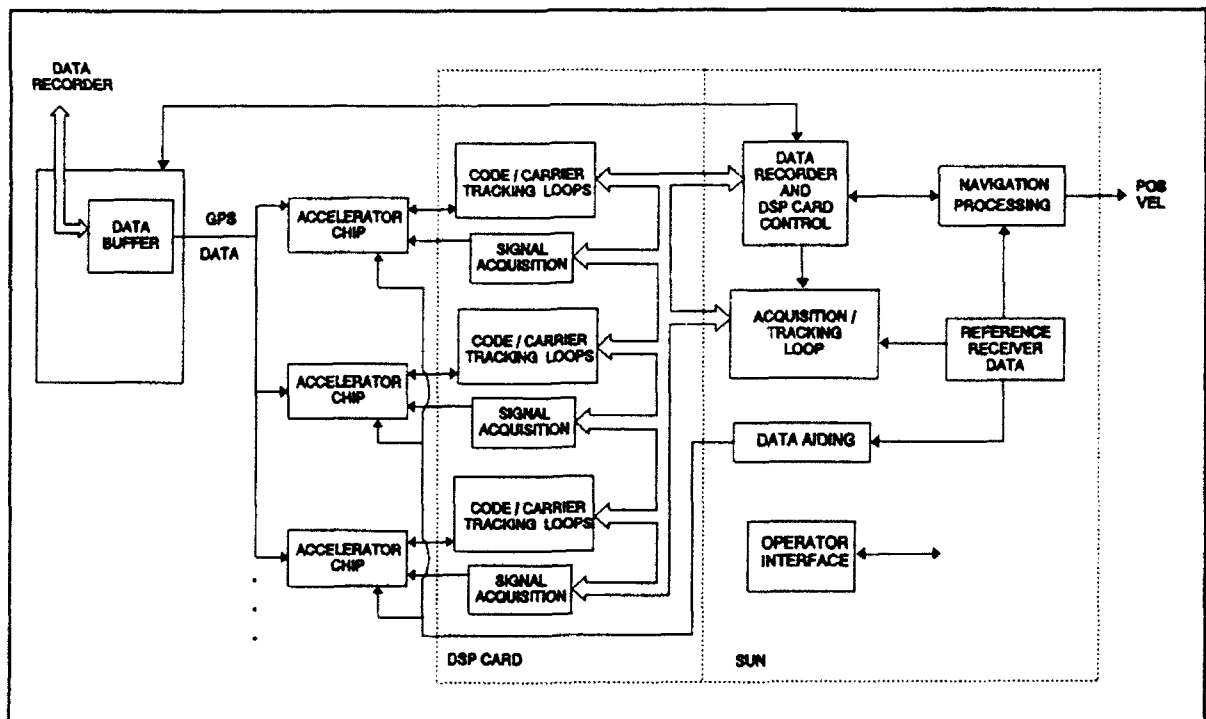


Figure 8 PTPS Processing Mode

## DISCUSSION

## E. SCHWEICHER

Instead of doing playback from your taperecorder, is it not possible to use Kalman filtering in case of break lock?

## AUTHOR'S REPLY

The benefit of playing back the data is that other sources of aiding data can be used. Kalman filtering is used anyhow.

# Communication Interference/Jamming and Propagation Analysis System and its Application to Radio Location

H. Kuzucu  
Communications Division  
SHAPE Technical Centre  
P.O. Box 174  
2501 CD The Hague  
The Netherlands

## SUMMARY

Modern defence systems depend on comprehensive surveillance capability. The ability to detect and locate the radio signals is a major element of a surveillance system. With the increasing need for more mobile surveillance systems in conjunction with the rapid deployment of forces and the advent of technology allowing more enhanced use of small aperture systems, tactical direction finding (DF) and radio location systems will have to be operated in diverse operational conditions. A quick assessment of the error levels expected and the evaluation of the reliability of the fixes on the targeted areas bears crucial importance to the effectiveness of the missions relying on DF data. This paper presents a sophisticated, graphics workstation based computer tool developed for the system level analysis of radio communication systems and describes its use in radio location applications for realizing such accurate and realistic assessments with substantial money and time savings.

## 1. INTRODUCTION

Radio direction finding and location capability is a key component of a comprehensive surveillance system which can be used in a multitude of applications such as communication intelligence (COMINT), electronic support measures (ESM), force strength assessments, emitter homing and targeting, interference and friendly source location. The requirement for more mobile surveillance systems in conjunction with the rapid deployment of forces and the advent of technology allowing more enhanced use of small aperture systems forces tactical direction finding (DF) and radio location systems to operate in diverse operational conditions. The effect of varying terrain structure on the propagation path losses and the noise introduced by jamming significantly affect the performance of DF stations and the siting requirements. Therefore for the effectiveness of the missions using DF data it is important to have an accurate assessment of the bearing error (BE) levels expected and a realistic evaluation of the reliability of the fixes on the targeted areas. SHAPE Technical Centre's Communication Interference/Jamming and Propagation Analysis System (STC CIPAS) is a very user-friendly computer-based tool for conducting such realistic assessments in a very fast and cost-effective manner. This is accomplished by allowing the user to easily control all the relevant scenario parameters down to a very high level of detail through a multiple pull-down menu style interface. Different parameter combinations, system/equipment options and deployment schemes can be quickly evaluated and compared.

The propagation model TIREM (Terrain Integrated Rough Earth Model) integrated into STC CIPAS uses digitized terrain elevation data (DTED) to calculate accurate path losses across the real terrain profiles. The realism and accuracy thus achieved may be regarded as the stage prior to actual field measurements which would be much more expensive and take longer.

Analysis has been confined to tactical communications in VHF and UHF bands. Uncorrelated noise jamming, propagation-induced errors and the thermal noise are the only error

sources considered. The effect of the propagation-induced errors on the rms bearing and position fixing is considered only in the sense of signal power attenuation. rms bearing errors are calculated as a function of the signal-to-noise ratio (SNR) as given for phase interferometer DF systems. Stansfield algorithm has been taken as the basis for the derivation of the radio location parameters.

## 2. ARCHITECTURE OF STC CIPAS

The foundation of STC CIPAS is the scenario development application. The scenario development application is used to construct communication scenarios that contain units (deployed assets) with specific equipments, on a background map. The background map is constructed using the DTED of the relevant area. The equipment includes communication devices and jammers. Performance parameters are entered by the user to define the capabilities of each type of equipment and these can be saved in a database of equipment types. The site-dependant parameters relating to the equipments are specified within relevant unit definitions. Selected sets of units can then be grouped together to partition a complex scenario and be saved as laydown files. Network files define the links between various participants. In the final phase of the scenario development, the map, laydown and network files are aggregated together to form the desired scenario. Scenario files can be saved and added to a database. The user can easily retrieve an available scenario file for using or editing purposes. Additionally, the scenario development application includes a number of map and display utilities and preferences features to help the user examine and tailor the created scenarios. Detailed description of all the utilities and features is given in Ref. 1.

The STC CIPAS has also an application component which allows the user to conduct detailed communication analysis on the constructed scenario. This component comprises three main utilities. These are site-to-site propagation analysis, interference analysis and communications coverage. With all three utilities, STC CIPAS gives the user the option of selecting from three propagation models. These models are TIREM of the US Department of Defense Electromagnetic Compatibility Analysis Centre (ECAC), free space path loss model, and radio line-of-sight (LOS) model. LOS model like TIREM uses the terrain profiles derived from DTED. TIREM operates in the 40 MHz to 20 GHz frequency band. It includes LOS, diffraction, atmospheric absorption, and tropospheric scatter loss algorithms. Ionospheric refraction has been excluded. For each mode of propagation, the model accommodates several sub-modules. Besides, various gap-filling interpolative sub-modules have also been provided in order to avoid wide swings and discrepancies at the transitions from one mode to another one. More detailed information on TIREM is given in Ref. 2.

The STC CIPAS allows for communication devices and jammers being able to operate using multiple carrier frequencies with variable passbands, and thus allows simulation of the effects of frequency hopping. Users can create complex antenna beam patterns and define passbands of communication devices and jammers in different ways. Relative antenna directivities, passband overlaps and receiver

front-end noise are automatically taken into account in all calculations.

The following sections will describe each of the three analytical tools mentioned above in conjunction with their application to DF and radio location examples.

### 3. SITE-TO-SITE PROPAGATION ANALYSIS

The site-to-site propagation analysis utility allows evaluation of path-loss between units defined in a scenario for which communication networks have been established. In addition to path loss, it calculates the received power levels, power and noise margins and jamming power if jamming is operative.

The sample scenario prepared for the demonstration of the site-to-site capability has been superimposed on a section of Germany as shown in Figure 1. Figure 2 shows the same section in full zoom. For the purpose of clarity, the simplest possible scenario involving all of the functionally necessary participants has been used. This consists of a mobile DF ground site (marked with BNFD - Battalion Fire Distribution Centre), a mobile ground transmitter (marked with BDEFDC - Brigade Fire Distribution Centre) which the DF site aims to locate and an airborne jammer which jams the DF site.

The DF site BNFD and BDEFDC are 52.5 km apart from each other. The airborne jammer cruises along a race track course at an altitude of 1000 m and aims to jam the DF receiver.

In the first example of the site-to-site analysis, the VHF transmitter on BDEFDC is tuned to 70 MHz with a bandwidth of 50 kHz as shown in Figure 3. It has an omnidirectional antenna and an effective radiated power (ERP) of 200 W. For the sake of performance testing, the DF receiver has been set to the same carrier with a minimum discernible signal (MDS) level of -115 dBm. A current technology phase interferometer DF station used for tactical intelligence collection would cover from 20 to 500 MHz with extendability to 1200 MHz. The IF bandwidth has been set to 50 kHz. The jammer parameters are set as shown in Figure 4. Jammer antenna is directed towards the DF station and has 0 dB gain.

With the values as given above, the site-to-site analysis is first carried out under no jam conditions using free space loss. The result is displayed in the first line of Figure 5. Free space loss is calculated as 103.6 dB. Since the threshold value of SNR for the DF is set to 0 dB (see Fig. 3) and jammer has not been turned on, the noise margin of 64.4 dB is the signal-to-noise (SNR) ratio calculated from the received signal (-50.6 dBm) and the receiver noise. The problem now is to translate the SNR into the rms bearing error of the DF station. In the following analysis, the observational errors and the environmental errors such as scattering, reradiation and propagation path deviations have not been considered. Effect of propagation induced errors has been accounted for only in terms of signal power attenuation.

As was mentioned earlier, the DF system used has been assumed to be a direct phase comparison interferometer. These systems are well suited for tactical intelligence collection and ESM applications. Both require near instantaneous DF over a broad range with precise angle-of-arrival (AOA) measurement accuracy. Further the signals of interest include a variety of modulation formats. The direct phase DF is relatively tolerant of signal modulation including short-duration and spread spectrum transmissions. Phase interferometers are subject to four major instrumental errors as discussed below (Ref. 3):

1. Antenna scattering and coupling;
2. Thermal Noise;
3. Multichannel phase mismatch and tracking;
4. Computational approximations.

The errors introduced by antenna scattering and coupling are deterministic; therefore these errors can be defined by site calibration and negated by the DF algorithm. Multichannel phase mismatch and mistracking errors are deterministic; however, these errors are a function of time, temperature, signal conditions and operational effects such as maintenance. Therefore periodic test signal insertion is required. Computational errors arise from the use of algorithms which approximate certain functions. Generally, algorithm errors are small if correction procedures are used and may be ignored if the correction schemes are robust. We will concentrate on the thermal noise and the uncorrelated noise jamming. The primary effect of these is to introduce random errors into the phase measurement process. Assuming a relatively high SNR, Figure 6 gives a plot of the root mean square (rms) bearing error versus SNR for a phase interferometer which has a baseline length to wavelength ratio of 0.5 (Ref. 3). In all the following calculations, zero mean rms bearing errors with Gaussian distribution have been assumed. Zero mean value allows to interchange the error variance with the square of the rms error. Where possible, the DF baselines are aligned with the coordinate axes such that the DOA is at 0 degree azimuthal and 90 degree elevation angles.

If we now apply the SNR value of 64.4 dB to the curve of Figure 6, the rms bearing error is found to converge to 0. The accuracy introduced through the use of TIREM is strikingly illustrated when the analysis is repeated with the same settings but using TIREM instead of free space loss. The second line of Figure 5 gives the TIREM results. The path loss prediction is 25.1 dB more than the free space value and the received power drops to -75.7 dBm. The SNR falls to 39.3 dB. Although significantly lower than the first value, this figure still yields a very low rms bearing error. The terrain profile between the DF site and the emitter is shown in Figure 7.

The next example further emphasizes the importance of realistic and accurate calculations by demonstrating the significant performance difference imposed by only a slight coordinate change on the emitter side of the link. Despite the coordinate change, the link length is almost the same. The new emitter coordinates and the new path profile is given in Figure 8. Comparison with Figure 7 indicates that the path is now obstructed with two knife edges rather than one and the first Fresnel zone covers more terrain. Consequently the received signal drops by more than 16 dB and the SNR is 23 dB. This is shown in the third line of Figure 5. That means that the rms bearing error would now be about 2 degrees.

If we now revert back to the original emitter coordinates but increase the frequency to 160 MHz, the SNR drops from 39.1 dB to 33.1 dB. This worsens the bearing error by 0.5 degrees.

Let us now turn the jamming on (see Fig. 4). The effect of this is illustrated comparatively in Figure 9. The second line in the figure indicates a received jamming power of -103.8 dBm. When compared with the unjammed value in the first line, this indicates a drop of 11.5 dB in the SNR. The bearing error rises to 1.2 degrees. The third line gives the result when the IF bandwidth of the DF receiver is increased from 50 kHz to 100 kHz. Both the receiver noise and the jamming noise double up reducing the SNR by half to 24.8 dB. This translates into an rms bearing error of 2 degrees.

Site-to-site analysis can also be applied to frequency hopping radios. In this mode, it gives an output averaged over the multiple frequencies in addition to the results for each individual frequency. The following is an example for the direction finding of a hopping emitter under full and partial jamming of the DF receiver. The multiple frequencies of the VHF emitter are chosen as shown in Figure 10. In the first run of the site-to-site analysis, 57 dBm jammer power directed at the DF site is split over five partial bands each of which is 100 kHz and centres one of the five emitter frequencies. Therefore each of the emitter frequencies received

at the DF site has jamming noise superimposed on it. The first six lines of Figure 11 gives the outcome for this case. The sixth line is the result averaged over the previous five individual values and corresponds to an rms bearing of 2 degrees. In the second run, jammer frequencies are offset such that only the first carrier frequency is jammed. The result is given in the second set of six lines in Figure 11. As can be seen, the average jamming power input the DF receiver is 7 dB lower than before and also lower than that capturing the first carrier frequency which is the only one jammed this time. This improves the rms bearing error to 1 degree.

So far we have considered the assessment of the expected rms bearing errors as a function of the scenario variables. It now remains to see what implications these constraints place on the problem of locating a radio emitter.

The most basic form of horizontal emitter location is to use the triangulation of the line-of-bearings (LOB) from three sites at known locations. In the presence of random errors, geometric estimation of position based on the error triangle contains deficiencies. Stansfield has developed a statistical method (Ref. 4) of position estimation which places emphasis on the random errors. Assuming a Gaussian distribution of the bearing errors characterized by the standard deviation or rms error, the positional uncertainty is designated by elliptical probability regions (PR) which also provide a measure of confidence of the positional fix.

In an active operational scenario, tactical DF information could be expected to cover a wide area of interest across the whole front of the friendly positions to an operationally useful depth in enemy-held ground. The size of the target area is related to the level of confidence that the target lies within it. For military purposes a confidence level of about 75% would be necessary if action were being taken on the information (Ref. 5). Therefore for a target area it is important to know the size of the uncertainty areas and the reliability of the fixes acquired from a DF system under various operational conditions.

As was shown previously, CIPAS provides an efficient means of incorporating the effects of jamming, receiver noise and propagation path losses into the rms bearing errors under diverse tactical scenarios. The following examples extend these ideas to illustrate the effect of jamming and propagation path losses on the position estimation.

The scenario layout is shown in Figure 12. Three mobile ground DF stations evenly spaced on a 50 km baseline provide a best position estimation (BPE) for an unknown target at 60 km away. Stansfield method has been used. As before, the DF stations are assumed to be phase interferometers having the SNR versus rms bearing error relation described by Figure 6. A stand-off airborne jammer tries to disrupt the DF stations. The DF stations have zero mean Gaussian distributed rms bearing errors. Figure 13 gives the site-to-site output with and without jamming. As can be seen from the first three lines, when the jammer is off the expected SNR from each DF is about 26 dB. This gives an rms bearing error of 2 degrees. Following the method of Stansfield, the area of the elliptical PR is found to be 26 km squares for a 75% probability. This is designated by the small ellipse in Figure 12. When the jammer is turned on the SNR values of the DF stations drop to about 16.5 dB yielding a bearing error of 5 for each station. This raises the area of PR to 162 km squares. This is marked by the larger outer circle in Figure 12. Therefore for a BPE at this certain point on the terrain, jammer causes the uncertainty area to increase more than 6 times.

In order to illustrate the accuracy provided by TIREM, let us now slightly change the position of the emitter keeping the distances the same as before. Jammer is turned off. Figure 14 gives the SNR values for the original and the new emitter positions. As can be seen from the second set of three lines for the new emitter position, the SNR value at

each DF drops to about 18 dB. This means an rms bearing error of 4 degrees. The PR area increases to a value of 104 km squares. A slight move of the emitter may introduce significant terrain obstruction into the DF path and may be as deleterious as jamming in terms of inflicting location uncertainties.

#### 4. COMMUNICATION COVERAGE

The communication coverage utility allows swift evaluation of the propagation coverage provided by a communication transmitter, with the aid of high-resolution graphics. Output from the utility is superimposed on the relevant terrain map as radials (radial coverage) emanating from the transmitter, or as cells filling a selected area (region coverage). Both are coloured according to a legend defined for the received power or SNR bands at the points of evaluation. The evaluation takes into account communication device parameters, site-dependant attributes, electromagnetic factors, and the effects of surrounding terrain elevation data on propagation path-loss. The effect of jamming from any jammer platforms that may be targeting the probing receiver can also be evaluated. The user can choose between fixed-frequency and frequency-hopping analysis. In the case of frequency-hopping analysis, coverage at each point is calculated on the basis of an average power or SNR value obtained over the frequencies defined for the communication device. Users can control many options to tailor coverage analysis to needs.

The sample scenario used for communication coverage has been superimposed on the section of Denmark shown in Figure 1. Figure 15 shows the selected section in full zoom. The participants of the scenario are a fixed ground site (marked with CRC - Control and Report Centre), and an airborne jammer. The aircraft flying over the land is an airborne DF station trying to locate the ground site. It has an altitude of 300 m AGL at 37.6 km from the CRC. It flies across a 100 km race-track course with a cross-range to range (CR-R) ratio of 2.7. Sparagna et al (Ref. 6) state that with minimum perpendicular error estimation algorithm of Stansfield accuracy is reduced at CR-R ratios above about 3 whereas a CR-R ratio below unity provides relatively large rms range errors. Therefore the optimum CR-R range is from about 1 to 3. The CRC has a UHF radio the coverage of which will be examined. The airborne jammer cruises along a race track course at an altitude of 1000 m over the sea and aims to jam the communication in an angular sector centring the DF aircraft. The distance of the jammer to the DF aircraft at the location shown in Figure 15 is 87.2 km.

For the example demonstrating the application of communication coverage to DF and radio location, the UHF emitter is tuned to 300 MHz. The antenna has a main beamwidth of 32 degrees, no side and back lobes. The ERP is 200 W. For testing purposes, the airborne DF receiver passband is set to 2 MHz centred around 300 MHz. As for the jammer, it has an ERP of 60 dBm which is spread over 2 MHz in 3 partial-bands. A band of 1 MHz centred around 299 MHz is the only one overlapping the passband of the airborne radio. This means that only 1/4 th of the total jammer power affects the DF receiver. Jammer antenna has been directed towards the DF aircraft with a main beamwidth of 20 degrees. The jammed sector is marked with the dashed lines in Figure 15.

With the values as given above, the coverage has first been done under jamming. Radial coverage has been used with TIREM at an angular resolution of 4 degrees. Probing DF receiver is airborne at an altitude of 300 m AGL. The cut-off value of the signal-to-noise (SNR) ratio is set to 5 dB in the colour legend. This means that the radials will be broken at points which have an SNR below 5 dB. This coverage diagram is shown in Figure 15. As can be seen, communication is partially denied in the jammed area and the average attainable communication range in the jammed sector is 28.5 km which falls short of the location of the DF receiver. That means that the SNR available at the airborne DF site is lower than 5 dB. This implies the bearings taken in the jammed sector will have very large rms errors and will be practically

useless in locating the ground emitter. The communication is resumed outside the jammed sector and stretches up to 110 km. Starting immediately from outside the jammed sector up to about 75 km, the SNR values stay between 20 and 40 dB allowing bearings to be taken with rms errors varying between 0.2 and 3.8 degrees. In order to assess the disruptive effect of the jammer, the analysis is repeated with the jammer disabled. This time the whole area up to 110 km is covered without any gaps. Besides the SNR values inside the jammed sector enhance significantly making the whole area up to 75 km suitable for reasonable bearing measurements.

## 5. INTERFERENCE ANALYSIS

The interference analysis application allows users to choose victim sites and communication devices and then select jammers or unintentional interferers that could disrupt the communications at the site. The application then calculates the interference power received from each interferer and the probability of interference in case of frequency-hopping devices. Evaluations can be performed for all four possible combinations of fixed frequency or frequency hopping interferer-victim pairs. Output information is both graphic and textual. In the context of DF applications interference analysis can be used for spectrum monitoring and interference source location as shown in the example below.

The sample scenario used for interference analysis has been superimposed on a section of Italy. The participants of the scenario are, a fixed ground site (CRC), a mobile ground site DF station (BNFDC) and an airborne jammer. The scenario window is shown in Figure 16. The distance between CRC and BNFDC is 25 km. The CRC is a friendly UHF radio site. The airborne jammer cruises along a short race track course at an altitude of 300 m and aims to jam the CRC radio. The distance of the jammer to the DF site is approximately 100 km.

The DF site aims to locate the position of the airborne jammer. Since the aircraft is flying relatively low, the depression angle is shallow. Therefore azimuthal DF may be applied. For the sake of demonstrating the interference analysis, the DF radio is tuned to 300 MHz with a fictitious passband of 4 MHz. The jammer is a partial-band jammer with an ERP of 60 dBm. The way the partial-bands are arranged is as follows: 500 kHz around 298, 305 and 310 MHz, 750 kHz around 299 MHz, 1000 kHz around 301 MHz and 1750 kHz around 302.475 MHz all totalling to 5 MHz. With these values, the total signal coming from the jammer is found to be -65.08 dBm. The receiver noise floor is -100 dBm due to the large passband. Therefore the SNR (the signal in this case is the jamming waveform) is 35 dB enabling bearing measurements at 0.5 degrees rms error. Figure 17 displays the textual output of the analysis. The figure of 4/6 shown under the Ratio column indicates that 4 out of the 6 partial-bands of the jammer hits the passband of the DF receiver. On the other hand, despite a distance of only 25 km the signal coming from the CRC radio is much smaller relative to the jamming signal. This is due to the high attenuation introduced by the terrain profile. In addition to the textual output, a spectral breakdown of the jamming signal power is provided in the frequency spectrum plot shown in Figure 18. Besides, colour coded arrows (see Fig. 16) are superimposed on the scenario which indicate the sources and the direction of the signals reaching the DF site.

## 6. CONCLUSIONS

Development of the STC CIPAS started in 1988. The aim was to produce a speedy and flexible tool to support the study of UHF electronic counter-countermeasures (ECCM) deployment (HaveQuick), system-level analysis and investigation of electromagnetic compatibility (EMC)/ECCM trade-offs. Starting from selected modules of the US command, control, communications and intelligence simulation (C3ISIM) software STC has developed a highly versatile radio propagation analysis tool, or CIPAS which can be effectively applied to DF and radio location applications. The extreme computa-

tional speed, accuracy provided by TIREM, user-friendliness, high resolution graphics capabilities and wide frequency range makes STC CIPAS suitable for realistic and cost-effective bearing error predictions, reliability assessments of positional fixes and radio location success rate evaluations in diverse tactical communication scenarios and radio location applications.

## 7. REFERENCES

1. Kuzucu H., "SHAPE Technical Centre Communications Interference/Jamming and Propagation Analysis System User's Manual", STC TN-408, October 1991 (NATO Unclassified).
2. Sciandra R., "TIREM/SEM Programmer's Reference Manual", ECAC-CR-90-039, July 1990.
3. Jenkins, H.H., "Small-Aperture Radio Direction Finding, Norwood MA: Artech House", 1985, pp. 143-147.
4. Stansfield, R.G., "Statistical Theory of DF Fixing", J.IEE (London), Vol.94, Part IIIA, No.15, March 1947, pp. 762-770.
5. Mansfield N.P., "Some Aspects of the Deployment and Performance of tactical VHF direction finding", STC TM-748, November 1984 (NATO Restricted).
6. Sparagna J., Oeh G., Huber A. and Bullock G., "Passive ECM : Emitter Location Techniques", Microwave J., May 1971.

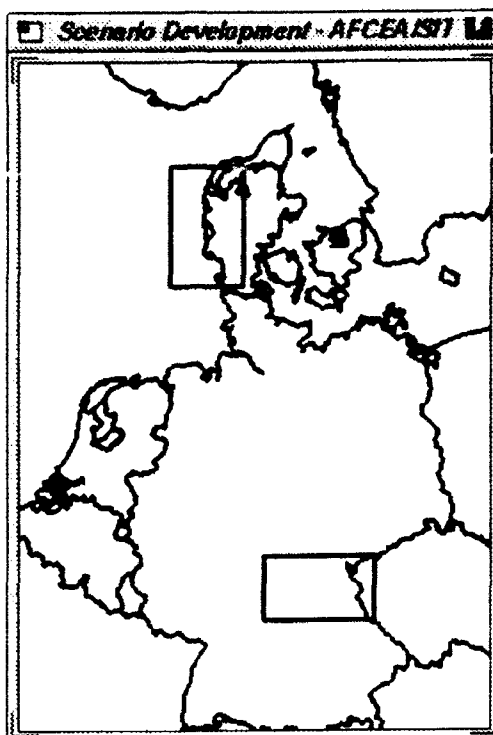


Figure 1 Scenario Locations

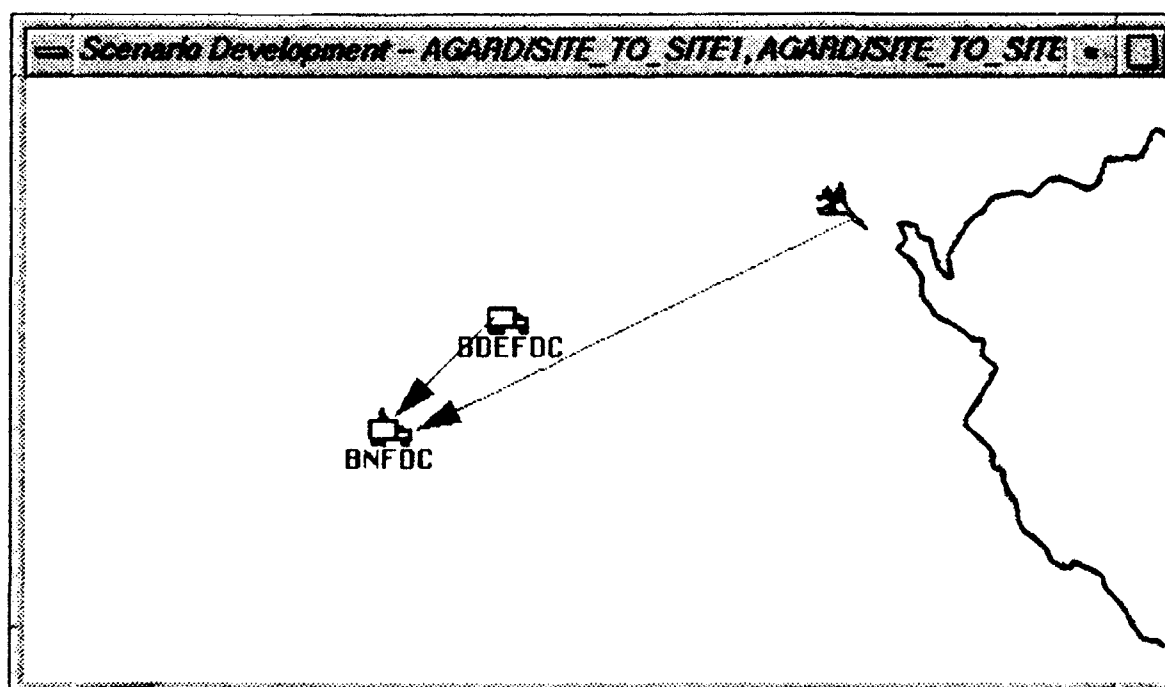


Figure 2 Site-to-site Analysis Scenario Window

**Communication Device Parameters:**

ID:

Antenna ID:

Power (dBm):

SNR (dB):

Protocol:

Noise Figure:

NDS (dBm):

☐ Set By Noise Figure  
☒ Set By NDS

**Auto Setting:**

Carrier Inc (MHz)

Number of Inc.

	Carrier (MHz)	PassBand (KHz)	
<input type="button" value="Add"/>	70.000	50.000	▲
<input type="button" value="Clear"/>			▼

Figure 3 Communication Device Definition Window

**Jammer Parameters:**

ID:

Antenna ID:

Power (dBm):

**Auto Setting:**

Carrier Inc (MHz)

Number of Inc.

	Carrier (MHz)	PassBand (KHz)	
<input type="button" value="Add"/>	70.000	500.000	▲
<input type="button" value="Clear"/>			▼

Figure 4 Jammer Device Definition Window



Power Margin Data:								
				Freq/Band	Loss	Rcv Power	Power Margins	Noise Margins
								Jammer Power
Site: BDEFDC_STE **Uni-Directional**								
Network: net1 Type: Duplex								
BDEFDC_STE	BDEFDC1	(C)	70/50	103.6	-50.6	64.4	64.4	N/A
Site: BDEFDC_STE **Uni-Directional**								
Network: net1 Type: Duplex								
BDEFDC_STE	BDEFDC1	(C)	70/50	128.7	-75.7	39.3	39.3	N/A
Site: BDEFDC_STE **Uni-Directional**								
Network: net1 Type: Duplex								
BDEFDC_STE	BDEFDC1	(C)	70/50	145.0	-92.0	23.0	23.0	N/A
Site: BDEFDC_STE **Uni-Directional**								
Network: net1 Type: Duplex								
BDEFDC_STE	BDEFDC1	(C)	160/50	134.9	-81.9	33.1	33.1	N/A

Figure 5 Site-to-site Analysis Output Window-Effect of Terrain on BE

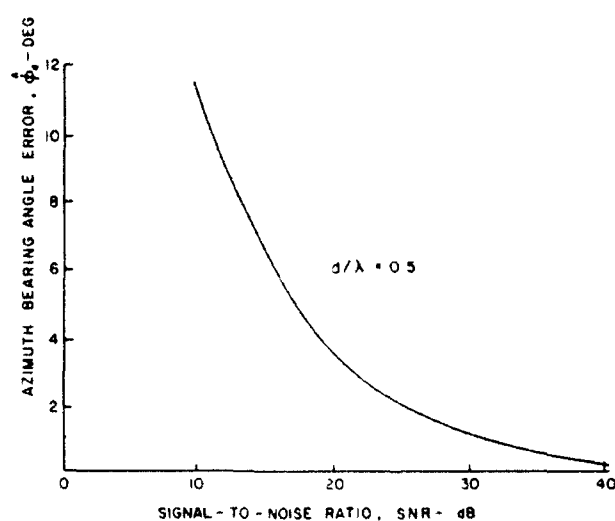


Figure 6 rms bearing error versus SNR - Phase Interferometer

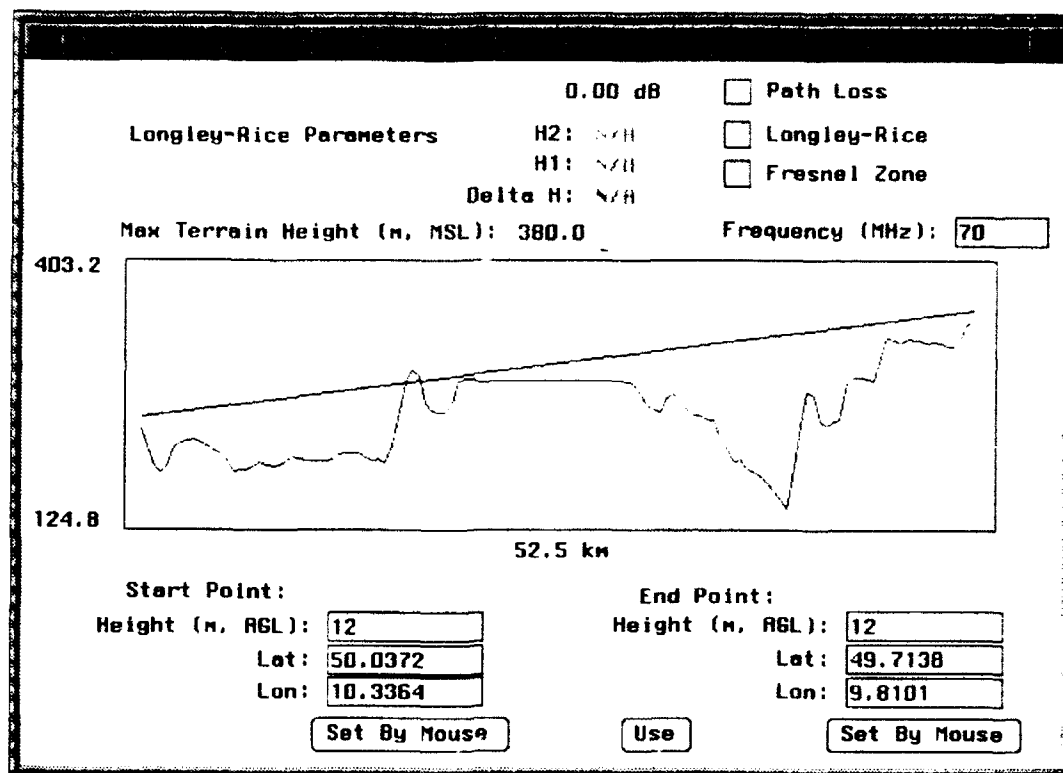


Figure 7 Side-view Profile - Initial Coordinates

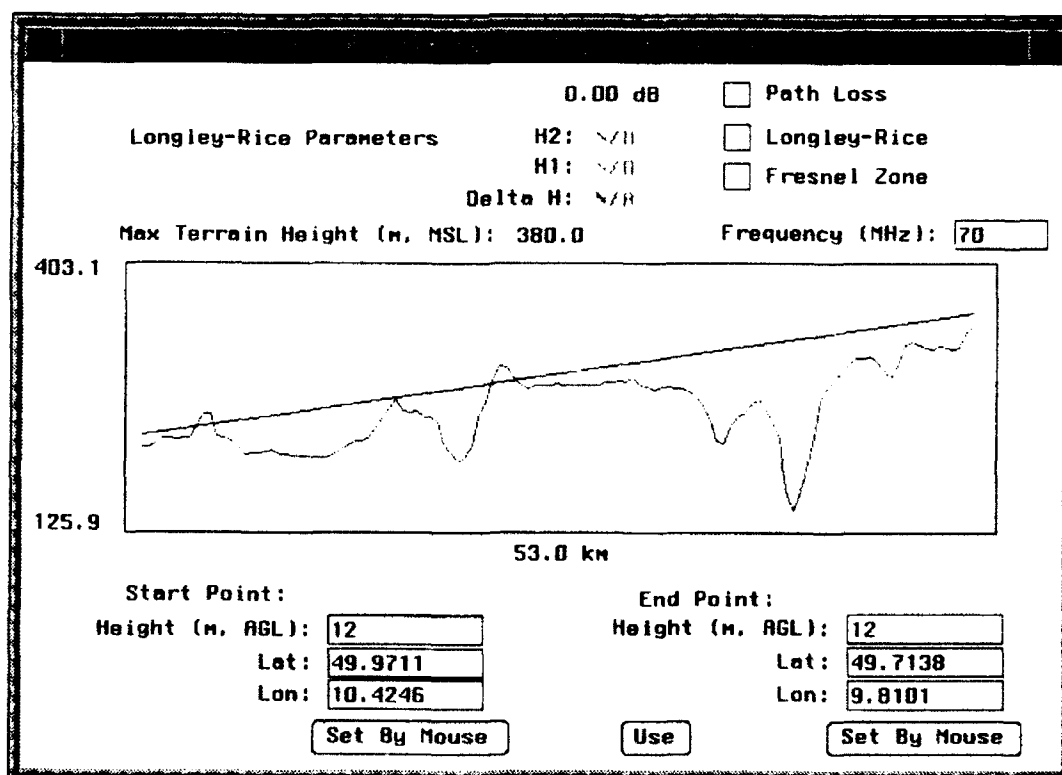


Figure 8 Side-view Profile - Modified Coordinates

Power Margin Data:

					Rcv Power	Power Margins	Noise Margins	Jammer Power
			Freq/Band	Loss				
Site: BDEFDC_STE **Uni-Directional**								
Network: net1 Type: Duplex								
BDEFDC_STE	BDEFDC1	(C)	70/50	128.7	-75.7	39.3	39.3	N/A
Site: BDEFDC_STE **Uni-Directional**								
Network: net1 Type: Duplex								
BDEFDC_STE	BDEFDC1	(C)	70/50	128.7	-75.7	39.3	27.8	-103.8
Site: BDEFDC_STE **Uni-Directional**								
Network: net1 Type: Duplex								
BDEFDC_STE	BDEFDC1	(C)	70/100	128.7	-75.7	36.3	24.8	-100.8

Figure 9 Site-to-site Analysis Output Window - Effect of Jamming on BE

Communication Device Parameters:

ID:

Antenna ID:

Power (dBm):

SNR (dB):

Protocol:

Noise Figure:

MDS (dBm):

☐ Set By Noise Figure ☒ Set By MDS

Auto Setting:

Carrier Inc (MHz)

Number of Inc.

	Carrier (MHz)	PassBand (KHz)
<input type="button" value="Add"/>	66.000	100.000
	68.000	100.000
<input type="button" value="Clear"/>	70.000	100.000
	72.000	100.000
	74.000	100.000

Figure 10 Communication Device Definition Window

Power Margin Data:

			Freq/Band	Loss	Rcv Power	Power Margins	Noise Margins	Jammer Power
Site: BDEFDC_STE "Uni-Directional"								
Network: net1 Type: Duplex								
BDEFDC_STE	BNFDC1	(C)	66/100	128.3	-75.3	36.7	24.6	-100.2
BDEFDC_STE	BNFDC1	(C)	68/100	128.5	-75.5	36.5	24.7	-100.5
BDEFDC_STE	BNFDC1	(C)	70/100	128.7	-75.7	36.3	24.8	-100.8
BDEFDC_STE	BNFDC1	(C)	72/100	128.9	-75.9	36.1	24.9	-101.1
BDEFDC_STE	BNFDC1	(C)	74/100	129.1	-76.1	35.9	25.0	-101.4
BDEFDC_STE	BNFDC1	(A)		128.7	-75.7	36.3	24.8	-100.8
Site: BDEFDC_STE "Uni-Directional"								
Network: net1 Type: Duplex								
BDEFDC_STE	BNFDC1	(C)	66/100	128.3	-75.3	36.7	24.6	-100.2
BDEFDC_STE	BNFDC1	(C)	69/100	128.5	-75.5	36.5	24.7	N/A
BDEFDC_STE	BNFDC1	(C)	70/100	128.7	-75.7	36.3	24.8	N/A
BDEFDC_STE	BNFDC1	(C)	72/100	128.9	-75.9	36.1	24.9	N/A
BDEFDC_STE	BNFDC1	(C)	74/100	129.1	-76.1	35.9	25.0	N/A
BDEFDC_STE	BNFDC1	(A)		128.7	-75.7	36.3	24.8	-107.2

Figure 11 Site-to-site Analysis Output Window - Frequency Hopping

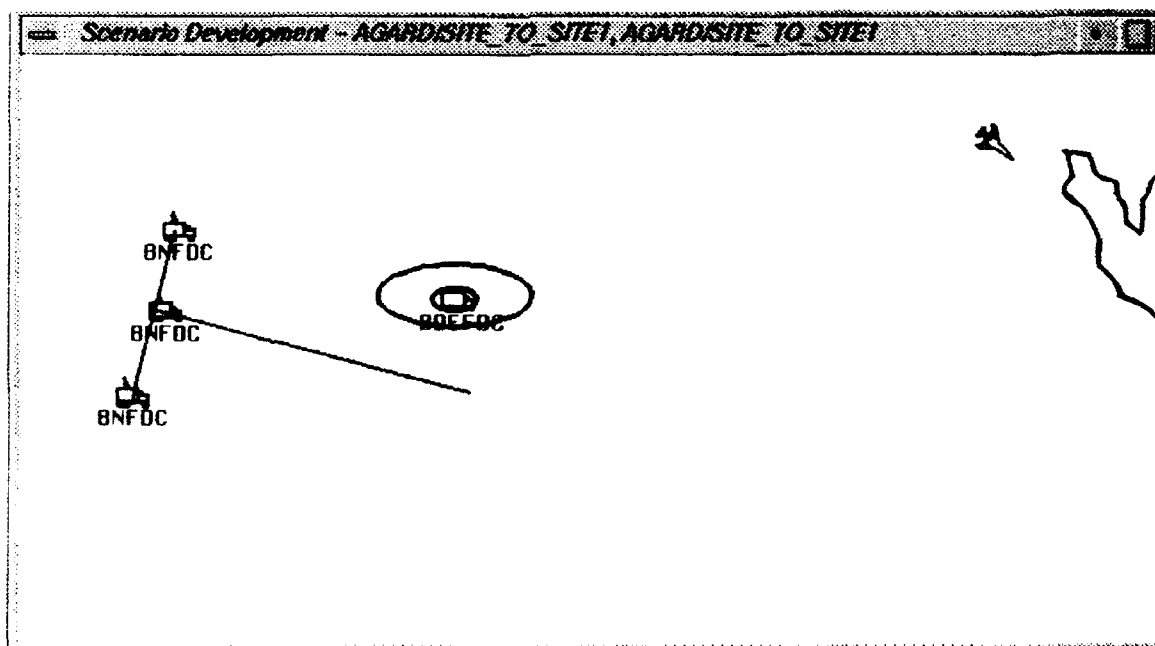


Figure 12 75% Probability Regions with and without Jamming

Power Margin Data:							
			Freq/Band	Loss	Rcv Power	Power Margins	Noise Margins Jammer Power
Site: BDEFDC_STE **Uni-Directional**							
Network: net1 Type: Duplex							
BDEFDC_STE	BNTDC1	(C)	160/100	138.4	-85.4	26.6	26.6 N/A
Network: net2 Type: Duplex							
BDEFDC_STE	BNTDC2	(C)	160/100	138.9	-85.9	26.1	26.1 N/A
Network: net3 Type: Duplex							
BDEFDC_STE	BNTDC3	(C)	160/100	137.8	-84.8	27.2	27.2 N/A
Site: BDEFDC_STE **Uni-Directional**							
Network: net1 Type: Duplex							
BDEFDC_STE	BNTDC1	(C)	160/100	138.4	-85.4	26.6	16.5 -102.3
Network: net2 Type: Duplex							
BDEFDC_STE	BNTDC2	(C)	160/100	138.9	-85.9	26.1	16.5 -102.9
Network: net3 Type: Duplex							
BDEFDC_STE	BNTDC3	(C)	160/100	137.8	-84.8	27.2	17.9 -103.2

Figure 13 Site-to-site Analysis Output window - Effect of Jamming on PR

Power Margin Data:							
			Freq/Band	Loss	Rcv Power	Power Margins	Noise Margins Jammer Power
Site: BDEFDC_STE **Uni-Directional**							
Network: net1 Type: Duplex							
BDEFDC_STE	BNTDC1	(C)	160/100	138.4	-85.4	26.6	26.6 N/A
Network: net2 Type: Duplex							
BDEFDC_STE	BNTDC2	(C)	160/100	138.9	-85.9	26.1	26.1 N/A
Network: net3 Type: Duplex							
BDEFDC_STE	BNTDC3	(C)	160/100	137.8	-84.8	27.2	27.2 N/A
Site: BDEFDC_STE **Uni-Directional**							
Network: net1 Type: Duplex							
BDEFDC_STE	BNTDC1	(C)	160/100	146.4	-93.4	18.6	18.6 N/A
Network: net2 Type: Duplex							
BDEFDC_STE	BNTDC2	(C)	160/100	146.3	-93.3	18.7	18.7 N/A
Network: net3 Type: Duplex							
BDEFDC_STE	BNTDC3	(C)	160/100	147.6	-94.6	17.3	17.3 N/A

Figure 14 Site-to-site Analysis Output window - Effect of Terrain on PR

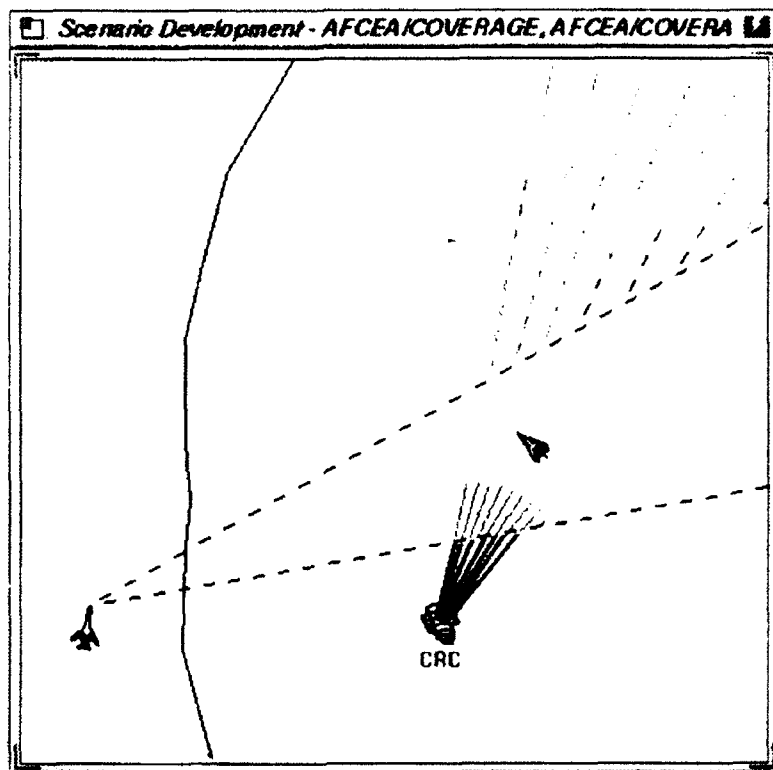


Figure 15 Communication Coverage under Jamming

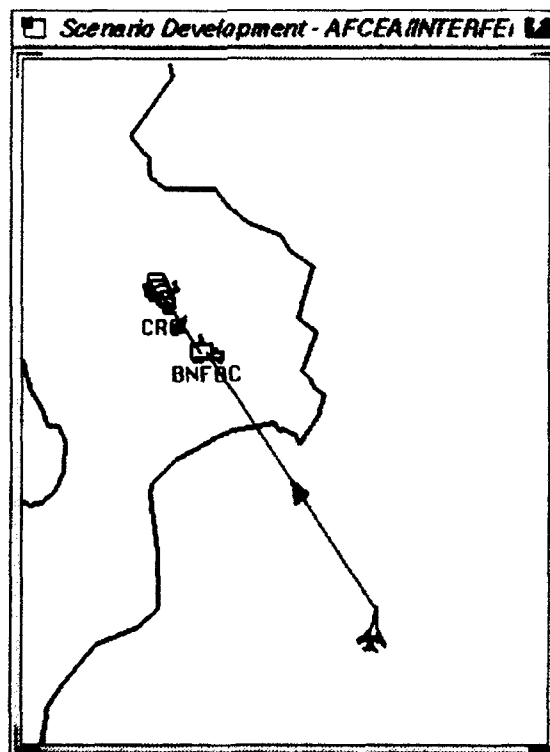


Figure 16 Interference Analysis Scenario Window

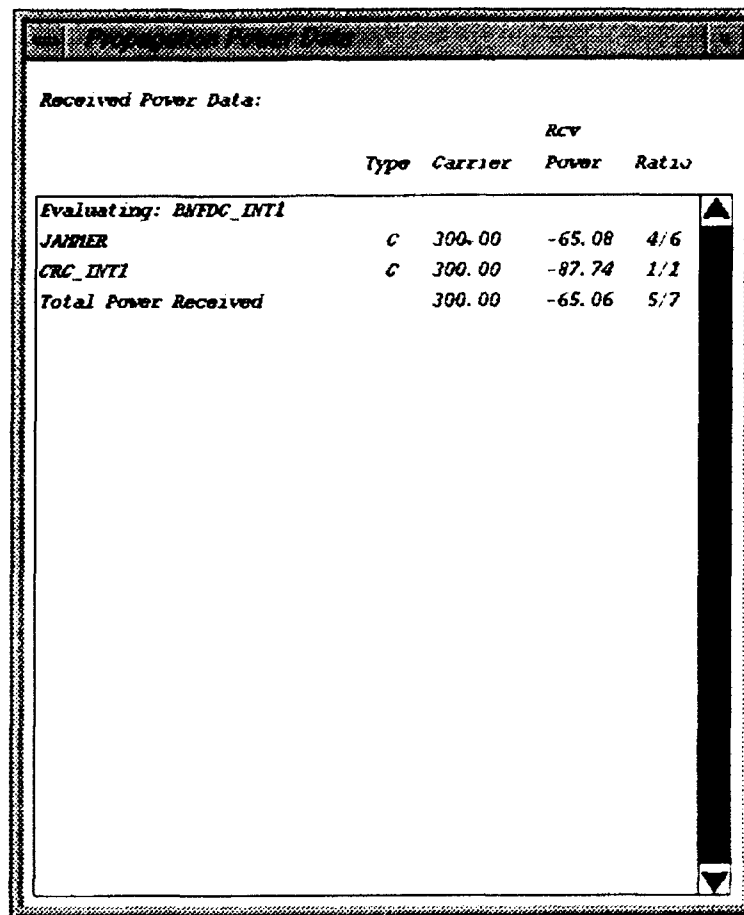


Figure 17 Interference Analysis Textual Output Window

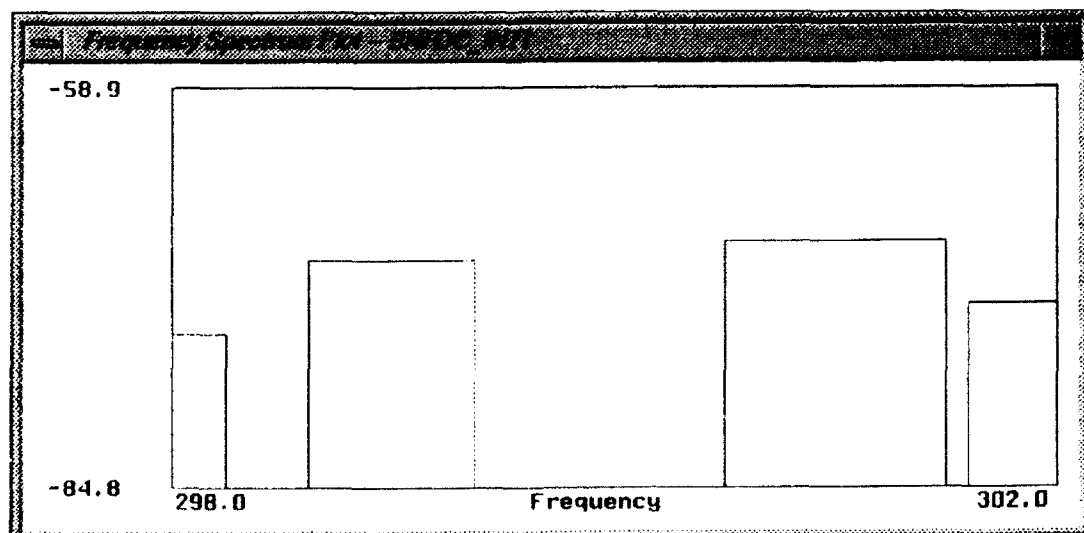


Figure 18 Interference Analysis Frequency Spectrum Plot

## DISCUSSION

### J. RICHTER

You stated that TIREM gives "accurate" propagation data over variable terrain and showed signal-to-noise and jamming diagrams. Since TIREM does not consider propagation anomalies and does not handle propagation over variable terrain other than diffraction, this conclusion is clearly not true. A scenario which is subject to ducting may encounter propagation conditions completely different from those predicted by TIREM. The user of your system should be aware of these limitations.

### AUTHOR'S REPLY

As I discussed in my presentation, TIREM covers LOS, diffraction and troposcatter modes of propagation. There are also 3 different LOS models depending on the minimum ray clearance above the terrain and the diameter of the first Fresnel zone. CIPAS as a whole used TIREM only as a sub-module. If under certain conditions, other propagation effects are to be taken into account that can easily be done inserting another module into CIPAS without changing any of the presented capabilities of CIPAS as a system level radio communication analysis tool. If you are aware of a better model than TIREM, I would certainly appreciate your contribution in the respect.

### G. HAGN

TIREM computes basic transmission loss, and the user must supply the internal power gain in the direction of the path. Currently available models for antenna gain (e.g., the Numerical Electromagnetic Code, NEC) assumes flat ground, but actual gains of antennas in irregular terrain can be different by 15 to 20 dB at low elevation angles (e.g., 0 to 5 degrees). It has only recently become possible to extend NEC to the case of irregular terrain on which the antenna is deployed. The strategy is to assume that the ground is flat within several wavelengths (i.e., locally flat) and use NEC to determine the currents on the antenna. These currents (and related fields) are then used as inputs to the NEC basic scattering code (NECBSC) developed at Ohio State University. The terrain is modeled as a lossy dielectric plate, and as many plates are used as are required to reduce the RMS difference between actual and modeled terrain to an acceptable value. This method was used to compute the perturbations in patterns (at 27.7 MHz) of 16 foot monopoles and  $\lambda/2$  dipoles at 15 feet above ground when located on the top of a hill, behind the hill and in front of the hill. Measured ground constants were necessary for the vertically polarized monopole, but perfectly conductive plates were adequate for the horizontally polarized dipole. The model work was performed by Professor J.K. Breakall, Mr. J.S. Young, and Professor D.H. Werner of Penn State University. This work is summarized in the proceedings of the Tactical Communications Conference (TCC-92), Fort Wayne, IN, April 1992: G.H. Hagn, J.K. Breakall, D.H. Werner, D.L. Faust, and R.W. Adler, "Measurement and Modeling of HF Monopoles and Dipoles in Irregular Terrain." It would be possible to generate a look-up table of gain changes for categories of antenna siting in irregular terrain. Such an addition could improve the realism of your simulations.

### AUTHOR'S REPLY

In CIPAS basic antenna patterns and antenna pair values can be simulated. If this information can be provided off-line, it is a matter of the user incorporating that data into the proper antenna elements. CIPAS capabilities would in that case be no different than that discussed in the presentation. In that respect I appreciate your comments on modeling the antenna gains and I will certainly consider the contribution of the work you are referring to.

### A. VAN DEN ENDE

What do you think about adding an estimate of the increase in rms bearing error in CIPAS in case of irregular terrain? If multipath effects are included, the presented rms bearing error is expected to be more realistic for a given terrain profile.

### AUTHOR'S REPLY

TIREM is only a sub-module of CIPAS. Any model, if it exists, which can give more accurate results than TIREM can be embedded into CIPAS without restraining any of its capabilities as discussed in the presentation. If you are aware of a better propagation model, I would certainly be interested in considering it. Your contribution in that respect would be highly appreciated.



# AIRBORNE SYSTEM FOR DETECTION AND LOCATION OF RADIO INTERFERENCE SOURCES

Bruno AUDONE - Alberto PASTORE

ALENIA - Settore Sistemi Difesa  
Stabilimento di Caselle  
Caselle Torinese  
10072 ITALY

## SUMMARY

The rapid expansion of telecommunication has practically saturated every band of Radio Frequency Spectrum; a similar expansion of electrical and electronic devices has affected all radio communications which are, in some way, influenced by a large amount of interferences, either intentionally or unintentionally produced.

Operational consequences of these interferences, particularly in the frequency channels used for aeronautical services, can be extremely dangerous, making mandatory a tight control of Electromagnetic Spectrum.

The present paper analyzes the requirements and the problems related to the surveillance, for civil application, of the Electromagnetic Spectrum between 20 and 1000 Mhz, with particular attention to the detection and location of radio interference sources; after a brief introduction and the indication of the advantages of an airborne versus ground installation, the airborne system designed by Alenia in cooperation with Italian Ministry of Post and Telecommunication, its practical implementation and the prototype installation on board of a small twin turboprops aircraft for experimentation purposes is presented.

The results of the flight tests are also analysed and discussed.

## 1. OPERATIONAL REQUIREMENTS

The rapid expansion of telecommunication has practically saturated every available radio frequency spectrum, making mandatory a tight control of it.

All radio communication transmissions are, in some way, affected by a large amount of interferences, either intentionally or unintentionally produced.

While the firsts are generated with the precise aim to disturb and to disrupt the communication channels and can be classified in the intentionally jamming category, either military for strategic purposes, or civil for criminal purposes, the second ones can be broadly subdivided in:

- spurious emissions and higher order harmonics originated by signals transmitted with not well controlled spectral bandwidth and frequency stability, at bands sometimes very far from victim frequency; these interferences are gene-

rally due to a bad design, manufacturing or maintenance of the transmitting station; this is particularly true for private radio and T.V. stations with high power and in proximity of victim frequency

- emissions generated by unauthorized transmitting stations (radio amateurs, citizen band, radiomobile services, etc.)

- emissions generated by scientific, industrial, medical, office equipment, domestic appliances (the microwave ovens are a good example) which do not comply with the interference suppression regulations

The operational consequence of all interference can vary by a very large amount, from simple, just annoying, flickering or white spots on T.V. receivers, to extremely dangerous, as the alteration or cancellation of communication, navigation or landing aids signals for civil and military aircraft, or disruption of military communication channels. As an example, a particularly dangerous situation occurred in Italy, at Rome Airport, when, some years ago, as reported by the press, a signal of unusual high power made completely useless six Air Traffic Control air-to-ground communication channels for over two hours; as a consequence, aircraft landings and take-offs were severely affected with very long delays.

A further recent entry in the already crowded R.F. spectrum is the Cellular Telephone service. Even if this service operates in a part of the spectrum which should be relatively free (in Italy the assigned frequencies span between 940 and 952 MHz) a potential hazard exists of reciprocal interference with other services.

The identification and location of all these interference sources is mandatory in order to pursue and to enforce the originators to stop the emissions or to reenter within prescribed limits and comply with internationally agreed regulations.

## 2. SEARCH FOR INTERFERENCE

The search for electromagnetic interference source shall be performed by means of the following activities:

- monitoring of the interested electromagnetic band
- analysis of the electromagnetic information

- location of the interference source
  - evaluation and comparison with applicable regulations
- The success of any action aimed to reduce or suppress the unwanted emissions rely primarily in the exact location of the emission source: the basis of the location techniques are:
- the identification of the direction of arrival of the signal, which, apart from propagation and measurements errors, is coincident with the direction of the transmitter
  - the accurate knowledge of the position from where the measurement of the direction is taken

To perform this task a radio direction finder shall be used, capable of operating over a wide frequency band, with any type of modulation and polarization, over full 360° in azimuth and 90° in elevation, giving accurate and reliable bearing even with low signal/noise ratio and short transmission period. This, of course, would be the ideal direction finder: a practical one will be a compromise among the abovesaid ideal requirements.

In Italy the Civil Authority designated for the control of the electromagnetic spectrum is the Ministry of Post and Telecommunication.

This control is presently performed with fixed installations and with mobile ground installations, the last used to survey the spectrum from several different positions and using direction finder equipment installed inside a van.

To increase the response time of the mobile installation, which is particularly critical when the duration of the interference is short, and to increase the radio horizon beyond the line of sight of a ground installation, Alenia, under the sponsorship of Italian Ministry of Post and Telecommunication, undertook the studies, developments, installation and tests of a Direction Finder in the band from 20 to 1000 MHz on board of an airborne platform.

The main purpose of this experimental installation and of the subsequent tests was not really an evaluation of the Direction Finder but a technical and operational evaluation of the whole system to gain an understanding of the problems connected with the airborne operation of a Radio Location System and to verify its suitability for a rapid location and identification of interference sources.

### 3. AIRBORNE RADIO LOCATION SYSTEM

The airborne installation is composed of two main elements:

- the Airborne Radio Location System
- the Aircraft Platform

They are presented in details in the following points.

#### 3.1. Airborne Radio Location System

The implementation used on the Alenia solution for the Airborne Radio Location System is formed by three main elements (Figure 1):

- the Direction Finder receiver
- the Aircraft Positioning System
- the Processing System

#### 3.1.1. Direction Finder receiver

The antenna is the key element of any Direction Finder. Because no antenna has a truly isotropic radiation pattern, any antenna has directional properties which can, in principle, be exploited as a direction finder.

In practice, however, there are some configurations and some techniques that are more suitable than others for direction finding purposes: to name just a few (the detailed explanation of which can be found in the text books) we can mention the simple vertical loop, the crossed loop, the Adcock, the interferometer technique, the Doppler, the phase comparison technique.

The Doppler is the technique utilized on Rohde & Schwarz PA-555, which was the Direction Finder used for Alenia experimental installation. A simplified block diagram is shown in Figure 2.

The receiving antenna is formed by a set of 16 vertical dipoles located on a circular pattern and electronically controlled in order to simulate two simultaneous clockwise and counterclockwise mechanical rotations. As a result, the simulated rotation of the dipoles within the electromagnetic field generates, due to Doppler effect, a sinusoidal variation of received frequency, increasing when the receiving antenna moves toward the incident wave and decreasing when it moves away.

The azimuth angle of the transmitter is obtained by comparison of the phase of the modulated signal with the phase of the signal that drives the antenna rotation.

The Doppler Direction Finder is particularly suitable for mobile applications because any Doppler effect produced by the speed of the vehicle where the antenna is mounted is automatically compensated.

There are three bearing outputs from the processing section of the Direction Finder:

- a digital output for a three digits indicator
- a digital output (IEEE 488 or RS 232) for connection to a data processing system
- an analog output for a CRT display: the analog presentation gives to the operator an indication of signal quality and an estimate of bearing accuracy.

The main characteristics of the PA-555 Direction Finder are the followings:

- Frequency Range : 20 - 1000 MHz
- Tuning : \* continuous with a knob  
\* discrete with a keyboard  
\* automatic with external computer
- Frequency accuracy :  $\pm 1 \times 10^{-8}$
- Minimum signal duration to enable D.F. :  $\geq 10$  ms
- D.F. Accuracy :  $\pm 2^\circ$  with the following minimum field strength:
  - \* 10  $\mu$ V/m @ 20 MHz
  - \* 3  $\mu$ V/m @ 30 - 300 MHz
  - \* 5  $\mu$ V/m @ 300 - 1000 MHz

Two different DF Antennas were used, covering respectively the frequency sub-bands 20 to 500 MHz and 500 to 1000 MHz.

### 3.1.2. Aircraft Positioning System

In order to establish the zone from where the R.F. signal is coming and, through triangulation from different subsequent positions, the location from where it is emitted, the bearings from Direction Finder must be correlated with the aircraft latitude, longitude and heading information.

To this purpose in the experimental installation a Loran-C receiver and a Gyroscopic magnetic compass was used. The accuracy of the position in geographic coordinates was in the order of 400 meters and the accuracy of magnetic heading in the order of 1°; these accuracies were considered sufficient for the initial operational assesment. The LORAN-C receiver caused many problems as in certain areas of the Country the coverage by Mediterranean Chain was marginal and the position data were often affected by unacceptable errors. A higher order of accuracy could be obtained, if required, using a GPS and an inertial navigator which imply, of course, an associated higher system cost.

### 3.1.3. Processing System

With aircraft latitude, longitude and magnetic heading available at the time when the bearing of the emitting source is taken from a minimum of two different points, sufficiently spaced apart, a simple trigonometric calculation can allow to identify the zone where the source is located (see Figure 3).

The uncertainty on the location of source is a function of the quantity of different bearings taken and of the errors on the data: for example, a 2° bearing error corresponds to 700 meters of across error in the position at a distance of 20 kilometers.

Clearly the accuracy is increased as more bearings are taken and when the different bearings are spaced between 60° and 120° apart.

For the experimental installation the calculations were performed at ground, either mathematically and graphically on a geographic map of the zone.

A data acquisition equipment was used which for every fix has stored and printed the following parameters:

- time
- frequency
- latitude
- longitude
- magnetic heading
- bearing

An sample of the printout is shown in Figure 4.

### 3.2. Aircraft

For the experimental installation an Alenia Partenavia P-68C-TC, two propellers, light aircraft was used (Figure 5).

The most significant characteristics for the specific mission are:

- Maximum Take Off Weight	1990 Kg
- Maximum Payload	690 Kg
- Cruising speed	240 Km/h
- Stall speed	140 Km/h
- Service ceiling	7000 m

The modification introduced on the basic aircraft to accommodate the Radio Location System were:

- Installation of a radome on the upper part of the aircraft center fuselage at a distance of about 30 centimeters from the metallic skin; the Direction Finder antenna is located inside the radome

- Installation of a Loran-C antenna on the upper part of the fuselage

- Installation in the passenger cabin of the following equipment:

- \* Direction Finder Antenna Scanning Unit
- \* Direction Finder Junction Box
- \* Direction Finder Receiver and D.F. Unit
- \* Loran-C Receiver Unit
- \* Data Acquisition Unit
- \* Interface Unit (to convert the analog output of gyrocompass to digital format of Data Acquisition Unit)

Figure 6 and Figure 7 show the details of the installation on the aircraft.

## 4. TEST RESULTS

The tests of the Airborne Radio Location System have been divided in three different series:

- ground tests
- airborne tests
- airborne Radio Location tests

The frequencies used for tests were selected in order to avoid interference: a transmitter at known location with the capability to vary the frequency was used, instead of Radio and TV broadcast station.

### 4.1. Ground tests

A series of tests were carried out at ground with engine running to give a first assessment of D.F. behaviour. During these tests, conducted over full frequency range from 20 to 1000 MHz and for complete 360° of azimuth, the average errors were contained within few degrees, even if for certain frequency below 100 MHz some points were in error on the order of about 10°.

### 4.2. Airborne tests

A first series of tests was carried out in order to establish the Direction Finder accuracy in operational conditions and at different frequencies over the full frequency range. The results are reported in table 1.

The errors were determined by the difference between the magnetic bearing,  $B_{mg}$ , defined as the angle between Magnetic North and the source measured clockwise from the receiver location, and the measured bearing,  $B_{me}$ , at various aircraft headings (Figure 8).

All measurements reported on table 1 were made at a di-

stance of about 50 km and at two different heights of 2500 and 5000 feet.

The aircraft headings were selected in accordance with paths permitted by Air Traffic Control authority.

A second series of tests was performed at different distance from the source; the maximum bearing errors are summarized in table 2.

Even if the errors found were larger than the results of laboratory tests in controlled conditions, they are however still suitable for the intended application.

An analysis of flight tests shows that:

- in proximity of the source the bearing indication change rapidly, probably due to diffractions produced by the aircraft metallic surfaces
- the bearings taken with source aligned with aircraft wings and tail are affected by the largest errors

#### 4.3. Emitter location tests

The main purpose of the evaluation of the system was to verify its capability to locate the emitting source.

To this purpose several flight tests were conducted using two different flight techniques:

- **homing**, where it is sufficient to maintain the aircraft heading aligned with the bearing until the bearing indication reverts to 180°, indicating the overflow of the source; this is the simplest and quickest method, but has some drawbacks, specially in presence of crosswind which forces to fly a logarithmic spiral path instead of a straight one
- **triangulation**, where the bearing of the source is taken from different positions, possibly separated in such a way to give bearings spaced between 60° and 120°

With the first technique all the sources used for tests were clearly identified, the only problem being the necessity to fly at low altitude to reduce the length of confusion cone over the source.

With the second technique the bearing data were plotted at ground: a sample of these plotting is shown in Figure 9. The bearings, taken at a distance from 6 to 10 km, gave the position of the source with an uncertainty in the order of 100 meters.

#### 5. CONCLUSIONS

An airborne radio location and monitoring system offers large possibilities for efficient and quick location of radio interference sources and for inspection of radio frequency spectrum.

Free line-of-sight and possibility to perform direction finding measurements from far places in a limited time shows that an airborne station can be efficiently integrated with ground fixed or mobile stations.

The experience made is considered positive and the possibility to carry out spectrum surveillance and to locate the emission sources over large areas in a very short time is not comparable with the capabilities of ground means used so far.

Geographical obstacles and difficulties in large urban areas due to reflections, usually encountered by ground means, are not present with an aircraft.

The use of an aircraft is also precious for rapid location of sources interfering with aeronautical communication and radionavigation services, not always possible with a ground mobile station which is limited to two dimensions only.

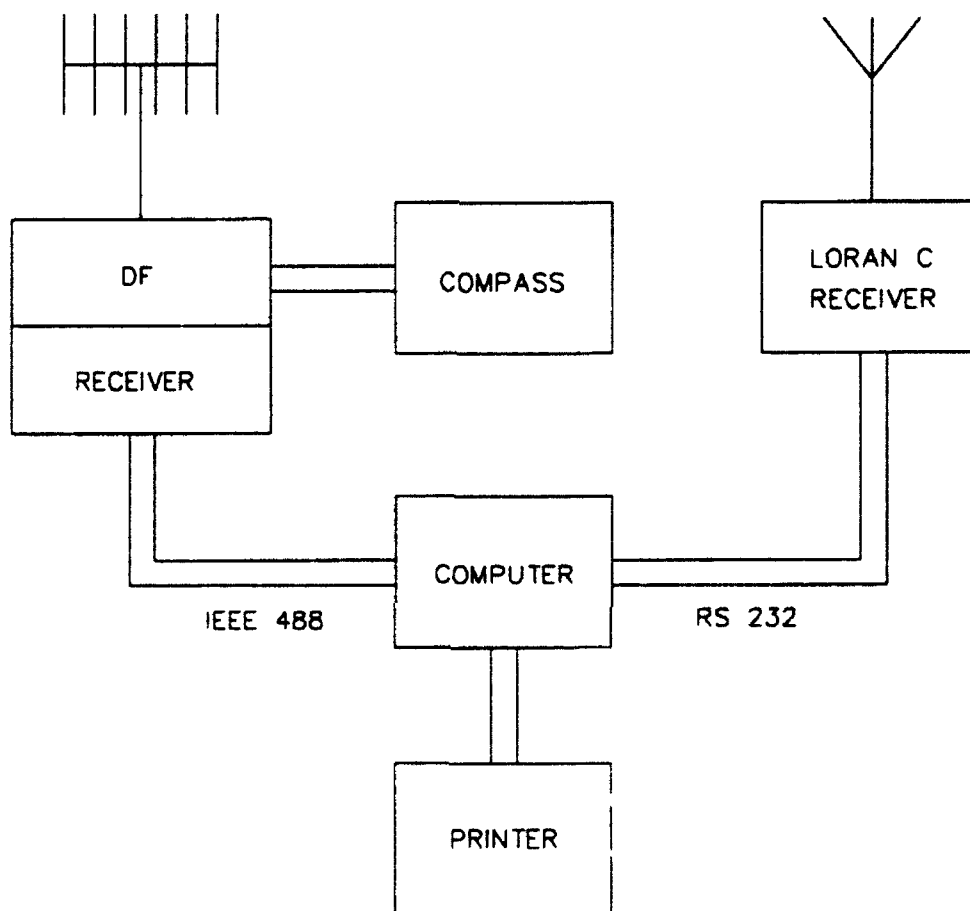


Figure 1 AIRBORNE RADIO LOCATION SYSTEM – BLOCK DIAGRAM

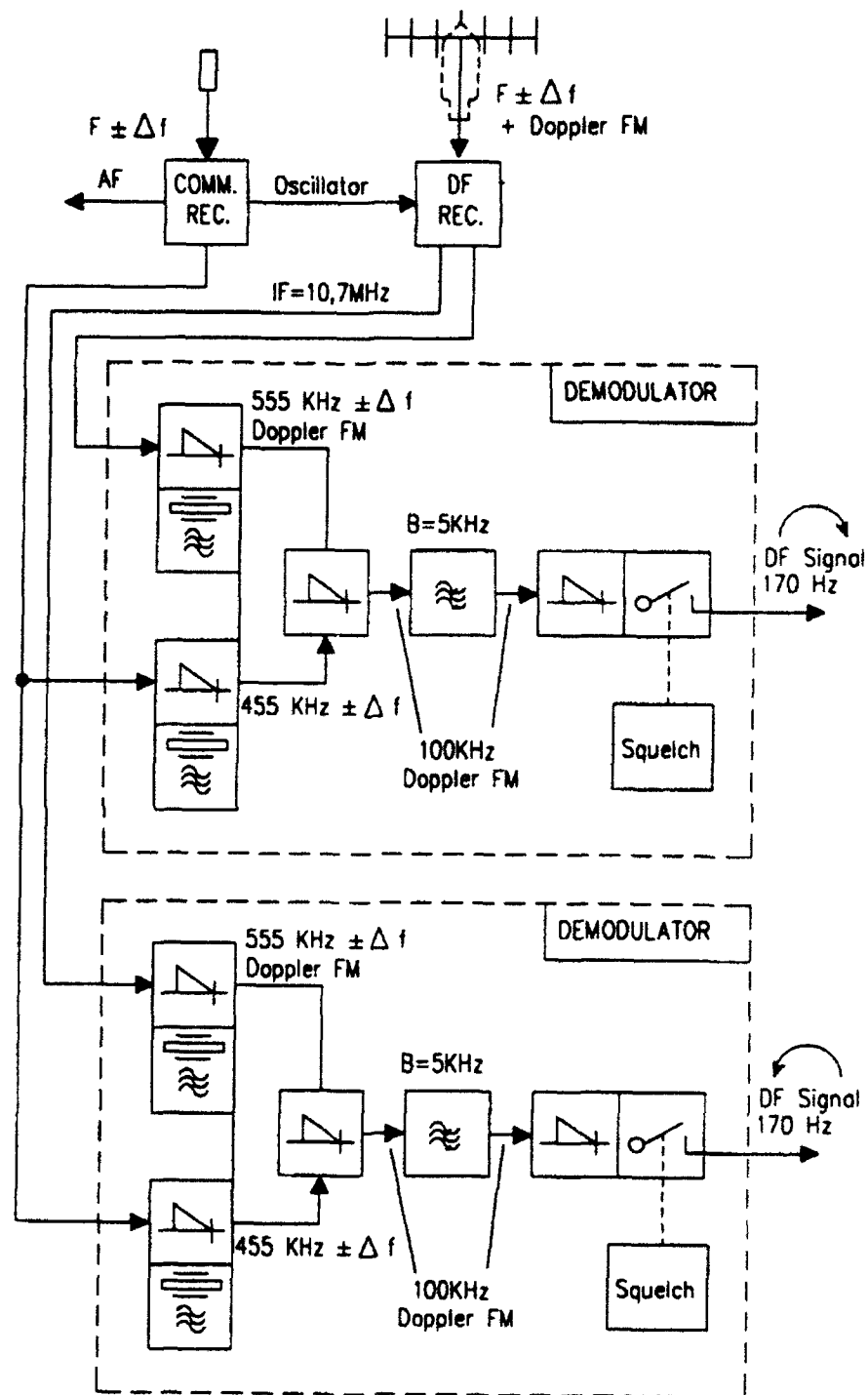


Figure 2 DIRECTION FINDER - BLOCK DIAGRAM

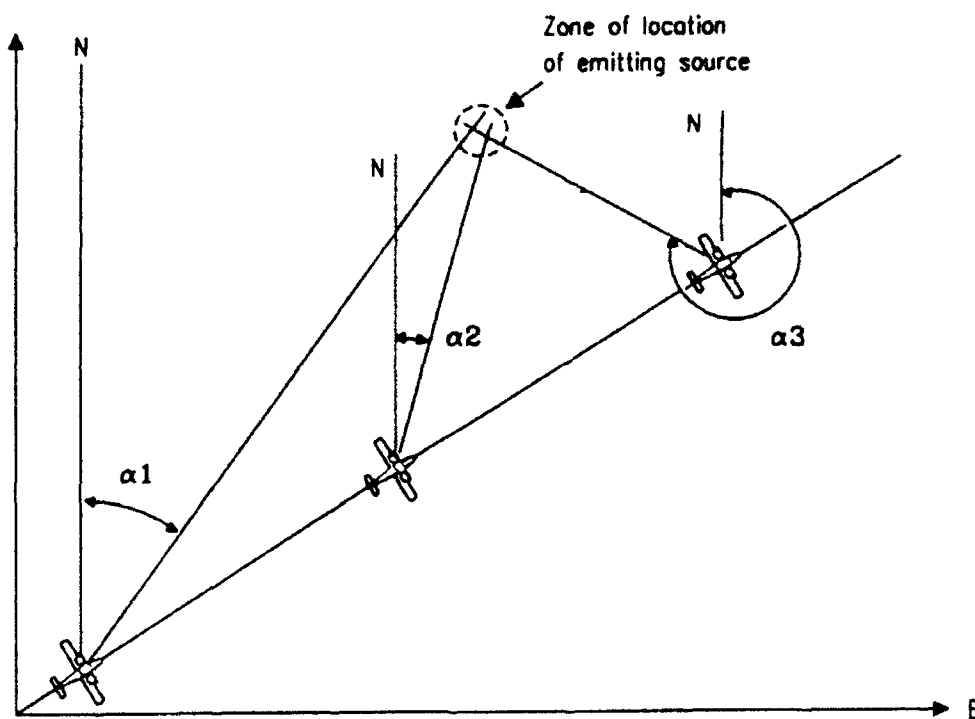


Figure 3 EMITTER SOURCE IDENTIFICATION

LATa=41.59,07		LONGa=12.34,27		FREQ=837.7 MHz	
LATp	LONGp	RILMAT	RILair	SCARTO	
42 13 4	12 51 67	223	224	-1	
42 13 5	12 51 71	223	227	-4	
42 13 7	12 51 75	223	227	-4	
42 13 8	12 51 80	223	227	-4	
42 13 9	12 51 84	223	227	-4	
42 13 10	12 51 88	223	227	-4	
42 13 11	12 51 92	223	227	-4	
42 13 11	12 51 96	223	227	-4	
42 13 12	12 52 0	222	227	-5	
42 13 13	12 52 4	222	227	-5	
42 13 13	12 52 8	222	227	-5	
42 13 14	12 52 13	223	227	-4	
42 13 15	12 52 17	223	227	-4	
42 13 15	12 52 21	223	227	-4	
42 13 16	12 52 25	223	227	-4	
42 13 17	12 52 29	223	228	-5	
42 13 18	12 52 34	223	228	-5	

Figure 4 SAMPLE OF THE PRINT-OUT

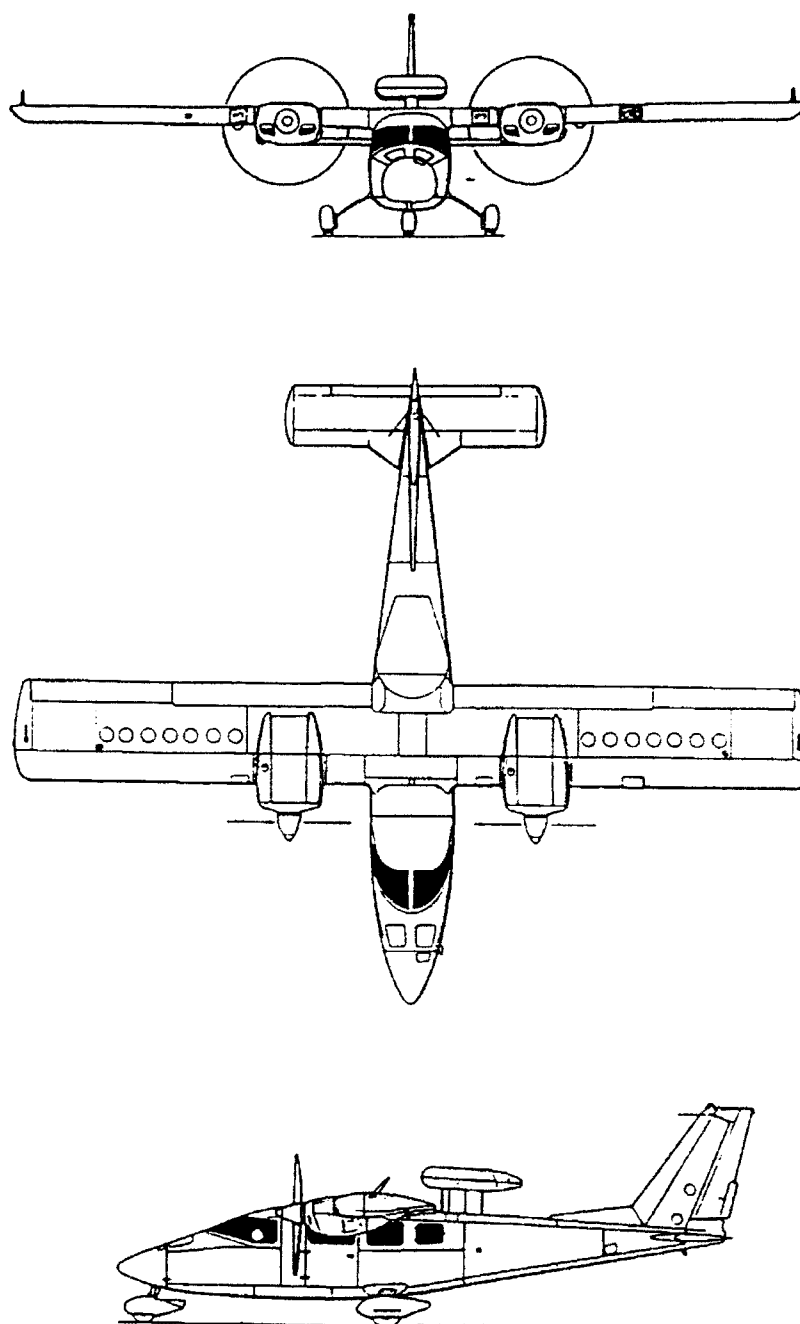


Figure 5 P 68 AIRCRAFT - THREE VIEW



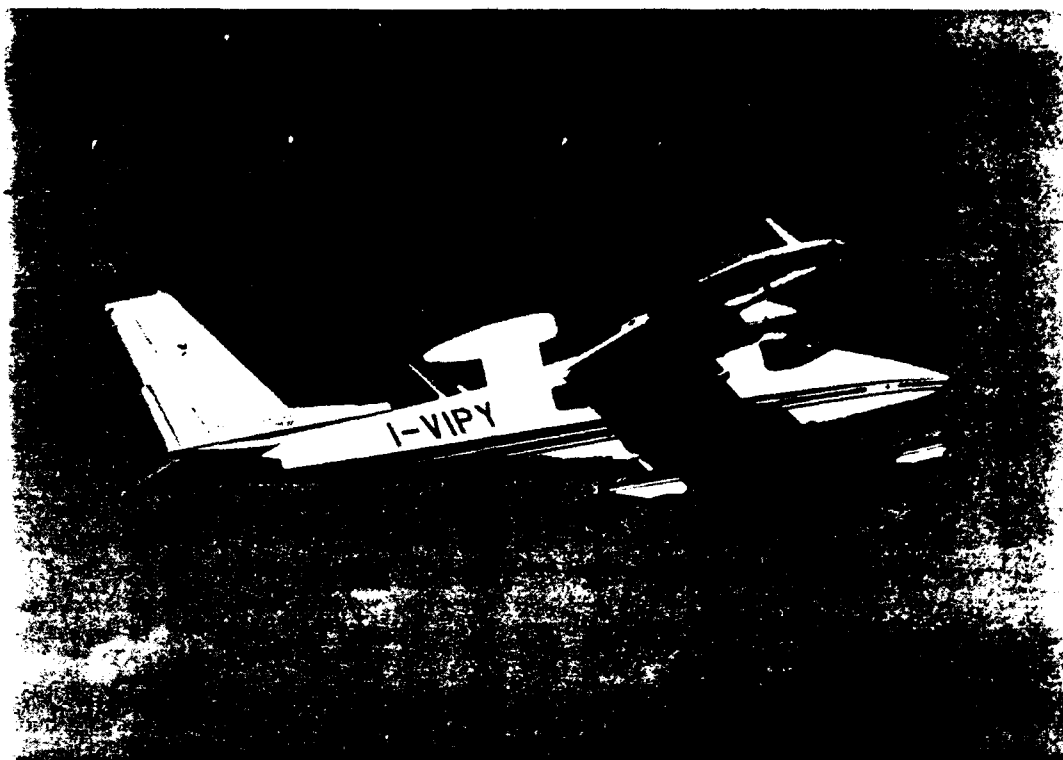


Figure 6 AIRCRAFT INSTALLATION – EXTERNAL VIEW

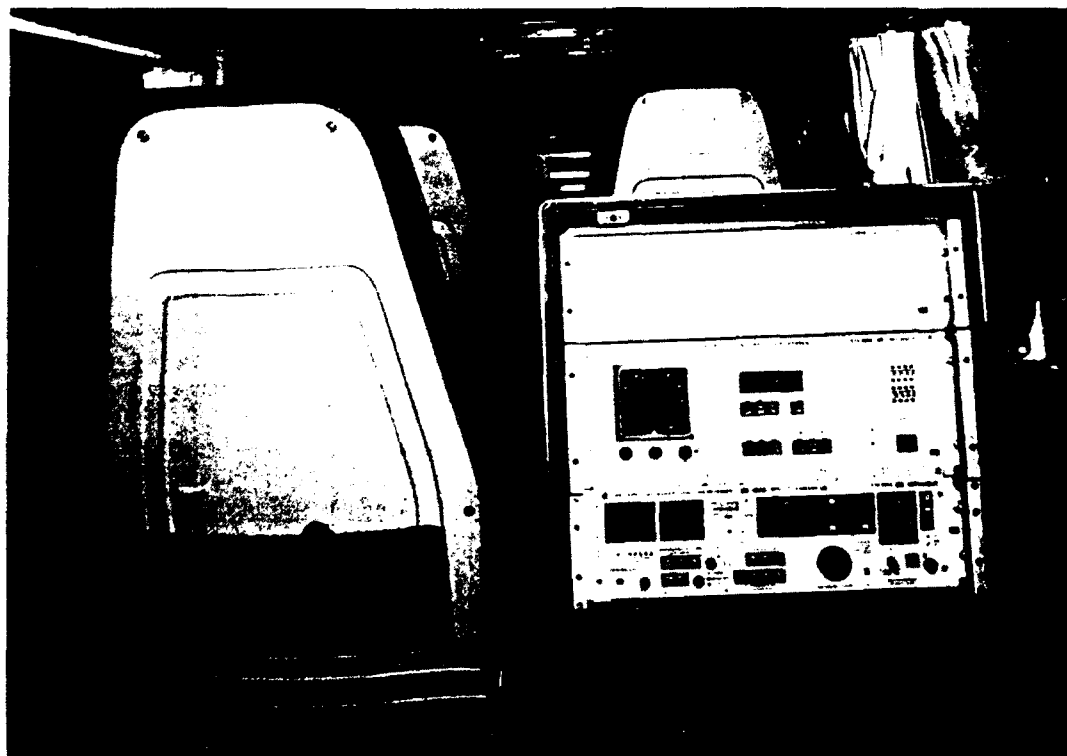


Figure 7 AIRCRAFT INSTALLATION – INTERNAL VIEW

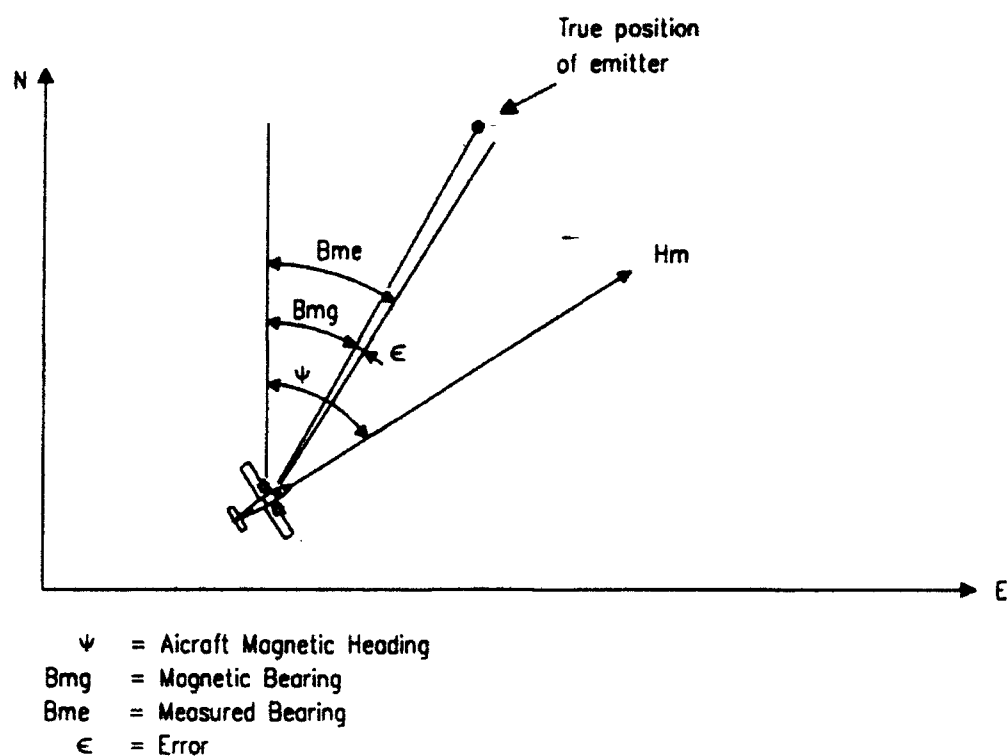


Figure 8 DETERMINATION OF ERRORS

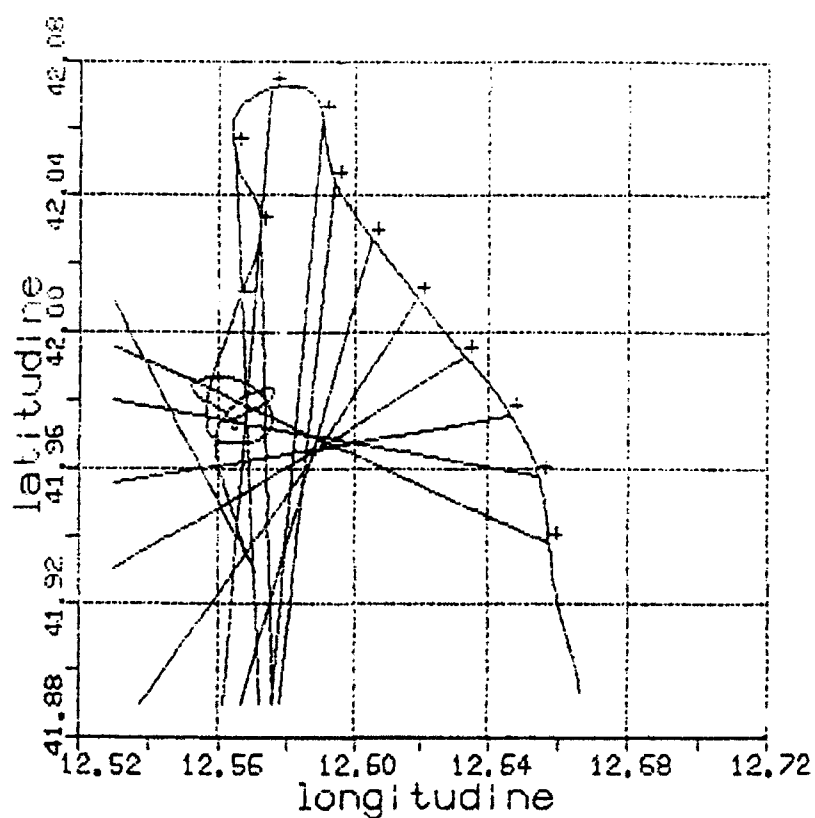


Figure 9 BEARINGS PLOTTING

FREQ.	AIRCRAFT	BEARING	AIRCRAFT	FREQ.	AIRCRAFT	BEARING	AIRCRAFT
(MHz)	HEADING	ERROR	HEIGHT	(MHz)	HEADING	ERROR	HEIGHT
	(deg)	(deg)	(feet)		(deg)	(deg)	(feet)
28.5	0	0	5000	132	0	0	5000
	25	+5			30	-1	
	90	+1			90	+8	
	115	0			115	+3	
	180	+4			180	-7	
	205	+2			205	0	
	270	+4			270	-5	
	295	+5			295	+5	
50	0	-4	5000	210	0	+4	2500
	25	+1			30	+5	
	90	-3			90	-5	
	115	+3			115	+2	
	180	-3			180	+1	
	205	0			205	+1	
	270	+3			270	0	
	295	+5			295	-1	
132	0	+3	2500	212	25	+5	5000
	30	+3			115	+2	
	90	+9			205	+1	
	115	-3			295	+3	
	180	-6		600	25	-3	2500
	205	+2			115	+1	
	270	+5			205	+5	
	295	+7			295	+5	
				900	115	+2	2500
					295	0	

TABLE 1

FREQ.	DISTANCE	MAXIMUM BEARING
(MHz)	Km)	ERROR
		(deg)
525	20	-5
615	160	+7
733	10	+8
805	80	-4
837	50	+7

TABLE 2

## DISCUSSION

### Y. TANIK

A simulated rotating antenna is used for direction finding, based on Doppler shift. This method has the disadvantage of being sensitive to modulation and multipath. Why did you prefer this method?

### AUTHOR'S REPLY

The Doppler effect produced by two different patterns rotating in opposite directions was preferred being more suitable for mobile application, particularly when the speed is high as in an aircraft, because Doppler effect caused by vehicle speed is automatically compensated.

### J. BENDER

Did you consider calibrating your DF-system as it is very sensitive to reflections from the platform?

### AUTHOR'S REPLY

The tests conducted so far were principally aimed to verify the operational implications more than the DF accuracy of the system. The possibility to incorporate in the processing unit a calibration curve for each frequency is presently under evaluation.

# LOCALIZATION SYSTEM BASED ON DATA MERGING OF IMAGES AND INERTIA.

J.Y. Cartoux, G. Sella, L. Agranier, T. Péchoux and G. Grenier.

Aérospatiale Missiles,

DE/ETRI

1 Rue Pablo Picasso,

78114 Magny-Les-Hameaux, France.

## Summary

The accuracy which is now required in the target approach of the modern missiles, requires the use of several sensors. Inertial systems, GPS when available, and imaging sensors are combined to improved the overall accuracy by the mean of filtering and data merging technics. Several tasks can be realized by the imaging system: targets detection and identification, missile position and motion estimation. The resulting informations will be used as command parameters under certain conditions of robustness and accuracy. Our goal is not only to compute these parameters but to specify their domain of usability. Our approach is based on the development of a global simulator. The models of the inertial systems, the informations computed by the imaging sensors are combined in a global Kalman filter scheme. The goal is to provide the user with the ability of studying the sensibility of the vision algorithms to some system parameters and external conditions in order to get the best overall accuracy.

In the first part of this paper, we present the Kalman filter scheme and the issues involved. In the second chapter, the data that can be obtained from imaging technics and the associated treatments are presented. We specify what are the constraints and the conditions under which these data can be obtained and study their sensibility to some system parameters.

## 1. Simulation scheme

Our interest is in the development of vision process which will be coupled to data coming from other different sensors in order to localize the vehicle or the objects seen by the imaging sensor with a better accuracy. The experimental conditions, the mission characteristics, the performances on an accuracy point of view, all of this will generate constraints on the choice of the sensors and on the vision algorithms. The purpose of the simulator we are working on, is to give to the user a global analysis tool for parametric studies in a development stage and for algorithms comparison and validation.

The simulation scheme is based on the Kalman filter concept. The filter seen as a merging filter has to estimate the parameters needed by the navigation process, by the guidance process or the target detection process. This implies to built the vehicle state and the relative target-missile state from all the informations coming from all the available sensors.

For the navigation problem, the vehicle state is defined by the motion of the vehicle mass center and Euler angles between the inertial reference frame and the vehicle

reference frame, and their derivatives. The guidance operation implies the knowledge of the relative motion between the vehicle and the targets. Furthermore, the quantities involved are different according to the chosen guidance law.

The classical optimal approach involves a Kalman filter which is based on a modelling approach:

- a modelling of the state evolution which in this case corresponds to the vehicle aerodynamic equations and to the kinematic behavior of a ponctual target in the guidance case;

- a modelling of the way by which the sensors measurements are linked to the state.

The mathematical formalism of the Kalman filter provides a theoretical optimality in the case of linear systems driven by Gaussian noises of null average and described by their covariance matrix. In fact, the real performances are strongly dependent on the modelling assumptions in the sense that the adequacy of the models to the real behaviors will induce the filter performance. Building such models implies a complete knowledge about the state evolution, the input forcing functions and the noise statistics but leads to very complex models which can be unusable in real-time applications with computer constraints. The solution is then to form simplified models. The development of a simulation tool will allow to examine the validity of the simplified models.

The structure of the global Kalman filter whose formalism can be found in [Bar89], is presented figure one next page. The building of the state vector is made during the initialization stage (figure 2) according to the available sensors. The state vector is defined by combination of elementary state vectors. The state, observation and covariance matrices are also built during this stage.

The sensors, when they are available, provide their measurements to the update part of the filter. Among them, the imaging system can provide several informations usable at different levels and for different tasks. Three different tasks can be realized: motion computation, scene structure reconstruction and target detection and localization. We will see in the next chapter what are the issues involved in these processes.

The control module is in charge of controlling the filter accuracy through the evolution of the covariance matrix. This matrix is a good description of the errors only if the models are good enough. If not, a high covariance will show the divergence of the filter, but the covariance can become very small. In that case, the gain is too small too, the filter considers that the estimation is correct and does not take into account much of the data. This can lead to the

situation where the estimates do not correspond to the real behavior even if the covariance matrix is very small. This can be avoided by controlling the parameters and the models. Different techniques are available to deal with non stationary models:

- non modeled perturbations can be introduced in noise terms: noise state and observation noise covariances are estimated on line from state update and innovation sequence [Mog89];

- impulsive perturbations can also be taken into account as inputs in the state equation whose unknown amplitude and time origine are estimated on line from innovation sequence according to a generalized likelihood criterium [Bas86];

- perturbations can be considered as modifications of some parameters in the model which can also be adaptively estimated [Bar89].

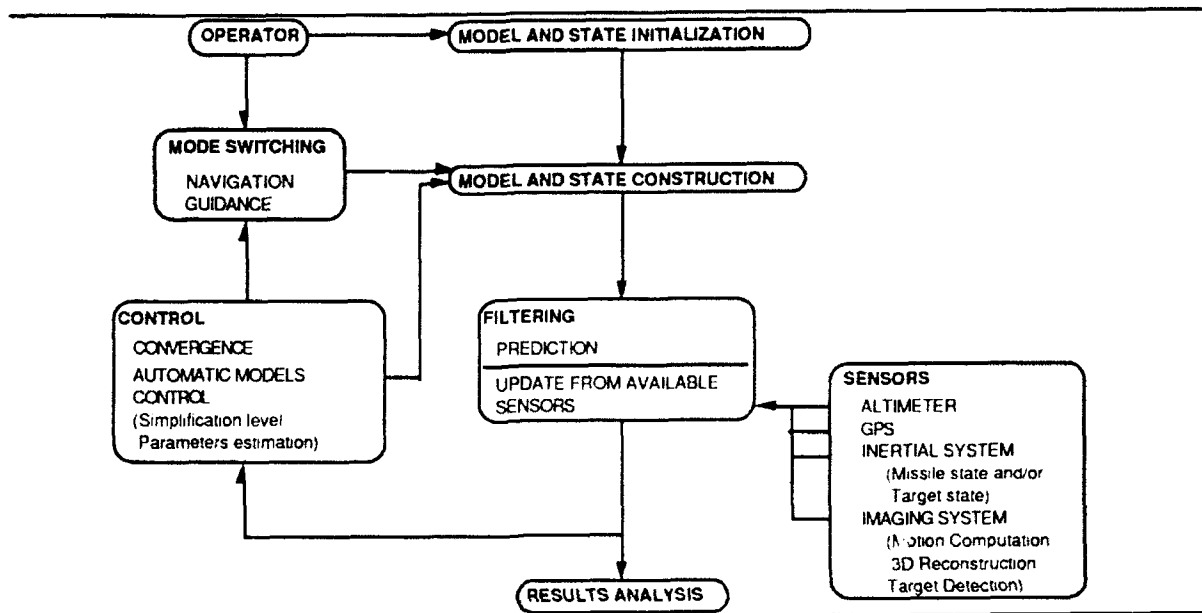


Figure 1.

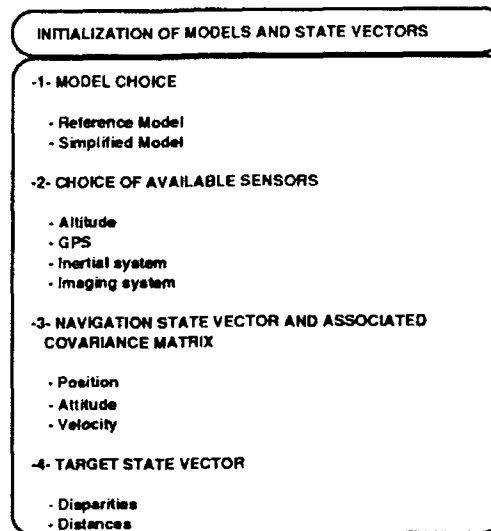


Figure 2.

## 2. Image algorithms and data

### 2.1 Motion.

Computing motion and structure from motion is a quite similar problem, the quantities and the methods involved in each process being quite identical. In both cases, the basic idea is to match characteristic points between images in order to compute either the motion of these points either, knowing the motion, their 3D position. These approaches are discrete approaches in the sense that the computation is made at regular time samples. They all made two basic assumptions: the environment is static and rigid, the observed terrain is flat. Three different types of approaches are mainly described in the literature: the techniques based on the optical flow or velocity field computation, the techniques involving the essential parameters computation based on the epipolar constraint, and the techniques based on the Focus of Expansion computation.

In the optical flow computation based approach ([Zin86],[Hag],[Ver89]), the estimation of the spatio-temporal derivatives of the intensity function leads to the velocity parameters estimation. The optical flow determination is not a simple problem and is generally quite expensive in terms of computer power.

The second approach is based on the computation of a particular point: the Focus of Expansion ([Rob91],[Bha89]), which is the intersection of the velocity vector with the image plane. The knowledge of this point leads to the determination of the motion parameters. The problem is that if the computation of this point for a translational motion is quite easy, the computation is not obvious for a translational and rotational motion and no methods have proved to be efficient.

The third approach is based on a simple equation which relies two points in the camera coordinate system at the two instants through the essential parameters matrix as they were called by Tsai [Tsa84]. This constraint simply express the fact that the two points, their projection in the image plane and the optical center positions are in the same plane, the epipolar plane. Several techniques ([Mai90]) have been used to minimize this constraint through a set of matched points. Generally, an inertial system is available onboard and can give a first estimation to the motion parameters: translation vector and rotation angles. This is one of the reasons for which we have chosen to iteratively estimate through a Kalman filter concept these parameters. Given the nature of the constraint, the quality of the estimate is strongly dependent of the stereoscopic basis. The accuracy will be acceptable when the displacement between the two images will be sufficient. The inertial sensors having a slow drift, the estimated values will then be used to correct the inertial system parameters or as input to the 3D structure computation algorithm presented below. One of the main issues for the validation of the process is to quantify the quality of the results in terms of the minimum number of points which is necessary and the sensibility to the terrain unflatness.

### 2.2 Scene Structure Reconstruction.

The goal of the reconstruction process is to estimate the position of ground points in the camera reference system. The knowledge of these informations allows either to compute a navigation information if the positions of the points in a geographic reference system is known, either to help the target detection and localization process, either to guide the missile towards the target. Among the methods available in the literature, we choose a Kalman filter based one [Jez90]. The points positions are estimated along time by the filter. The motion is supposed given with a covariance matrix associated and can either result from the other sensors or be computed by the method described above. The observation equation links the points coordinates in the 3D space to the image coordinates through perspective projection according to a classical pin-hole camera model. This equation is not linear and is linearized around the state estimate. The initial value of the state vector is computed from the informations given by the other sensors. The accuracy of these sensors allows the initialization of the covariance matrix.

At the current state of our studies, the goal is to measure the influence of some parameters on the accuracy of the estimation. The input of the algorithm is a sequence of matched points which are artificially generated data, the parameters being chosen for each test by the user. The points we are working on are:

- the compromise between the sampling time and the distance observability
- the influence of the precision of some system parameters like motion and velocity on the final estimations.

### 2.3. Targets detection

A target detection process is based on two steps which are extraction and filtering. The goal of the extraction process is to extract of the images, points or regions which are supposed to belong to targets. The criteria generally involved physical characteristics which are seen by the image sensor as specific gray level values like hot points for infra red sensors or polyhedra for radar sensors which both lead to high gray level values in the image. Due to the image formation process, noise is generated which produce false alarms in the detected targets list. The second stage consists then in filtering this list according to criteria specific of the targets like position, size, temperature even the ground repartition for targets formations. The knowledge on the observed scene of a distance information combined to these characteristics, will increase the selectivity of the filter. For an image sensor onboard of an air vehicle and looking toward the ground, this distance information is theoretically given by the ratio of the altitude of the vehicle by the tangent of the sensor inclination angle. These two parameters are known with a poor accuracy which makes the straightforward application of this simple formulae impossible, the targets filtering criteria becoming then unreliable.

The reliability of the filtering stage can be improved by the 3D reconstruction process described above. The tracking over time of the alarms and their position estimation by the Kalman filter provides a better accuracy for the distance estimation.

### 3. Conclusion

A simulation scheme for the localization of an air vehicle by combining inertial and imaging sensors has been presented. The simulator is based on a Kalman filter concept involving elementary state, observation and covariances corresponding to each sensor model and measurements set. Among these sensors, the imaging system can provide through different algorithms involving a Kalman filter, useful data for navigation, targets detection and guidance. The purpose of the simulator is to provide the user with a global analysis tool allowing the validation and the comparison of vision algorithms and parameters adjustment according to a predefined experimental environment including the mission type and the available sensors.

### References

- [Tsa84] R.Y. Tsai and T.S. Huang, Uniqueness and Estimation of Three-Dimensional Motion Parameters of Rigid Object with Curved Surfaces, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:13-27, 1984.
- [Maï90] H. Maître, Estimation du Mouvement d'une Caméra à partir de Deux Images Successives, *Technical Report, Ecole Nationale Supérieure des Télécommunications (ENST)*, Paris, France, May 1990.
- [Jez90] J.L. Jezouin and N. Ayache, 3D Structure from a Monocular Sequence of Images, *In Procs. of Int. Conf. on Computer Vision*, Osaka, Japan, December 1990.
- [Ver89] A. Verri and T. Poggio, Motion Field and Optical Flow, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:490-498, 1989.
- [Koch] R. Koch, R. Bader and W. Hinding, A Study of an Integrated Image and Inertial Sensor System, *AGARD-CP-474, Knowledge-based System Applications for Guidance and Control*.
- [Hag] E. Hagen and E. Heyerdahl, Navigation by Optical Flow.
- [Zin86] H. Zinner, Determining the kinematic parameters of a moving imaging sensor by processing spatial and temporal intensity changes, *J. of Opt. Soc. Am. A*, Vol. 3, N°9, September 1986.
- [Rob91] B. Roberts and M. Bazakos, INS Integrated Motion Analysis for Autonomous Vehicle Navigation, *Procs. of SPIE, Vol. 1521, Image Understanding for Aerospace Applications*, 1991.
- [Bha89] W. Burger and B. Bhanu, On Computing a Fuzzy Focus of Expansion for Autonomous Navigation, *in Procs. of Int. Conf. on Computer Vision*, 563-568, 1989.
- [Mog89] A. Moghzddznjoo and R.L. Kirlin, Robust Adaptive Kalman Filtering with Unknown Inputs, *IEEE ASSP Vol. 37, N°8*, August 89.
- [Bas86] M. Basseville and A. Benveniste, Detection of Abrupt Changes in Signals and Dynamic Systems, *LNCIS N°77*, Springer Verlag 1986.
- [Bar89] Y. Barshalom, T.S. Fortman, Tracking and Data Association, Academic Press, New York 1989.

### DISCUSSION

#### E. SCHWEICHER

1. Votre système est-il spécifique pour les centrales inertielle à composants liés (strapdown) ou peut-il être utilisé pour tous les types de centrales inertielle?
2. Avez-vous utilisé les fractales pour la création de vos images synthétiques?
1. *Is your system specific to inertial reference units with strapdown components or can it be used for all types of inertial units?*
2. *Have you used fractals for creation of you synthetic images?*

#### AUTHOR'S REPLY

1. Actuellement, le système est spécifique des centrales à composants liés.
2. Non.
1. *At present the system is specific to strapdown type inertial reference units.*
2. *No.*

#### G. HAGN

You mentioned image stabilization. Vehicles, such as helicopters hovering near the ground or land vehicles traversing irregular terrain, can cause relatively large amounts of platform motion. For how much motion can you still achieve image stabilization?

#### AUTHOR'S REPLY

The electronic stabilization system is a complement to a mechanical system. Therefore, it will correct for small angles and small amounts of vibration. This provides for very sensitive stabilization.



# THEORIE FRACTALE DES GRANDS RESEAUX D'ANTENNES LACUNAIRES

## FRactal Theory for Large Lacunar Antenna Arrays

C. GOUTELARD

LETTI  
Université Paris-Sud  
Bâtiment 214  
91405 Orsay Cedex  
France

### I. - INTRODUCTION -

La complexité des réseaux est directement liée au nombre d'aériens qui les composent.

Les lois de rayonnement en champ lointain sont bien connues et se déduisent de la loi d'illumination de l'antenne. Le souci du concepteur est, en permanence, de déterminer, pour un nombre d'aériens disponibles, leur répartition spatiale qui assure la loi de rayonnement la plus avantageuse. L'orientation électronique du lobe d'antenne et le fonctionnement en large bande rendent le problème plus complexe. Le souhait d'obtenir un lobe de rayonnement principal étroit et de limiter l'amplitude des lobes secondaires est toujours contradictoire.

Les structures régulières utilisées fixent les caractéristiques des diagrammes de rayonnement dans lesquels les variations qui obéissent à des lois connues sont peu favorables à l'optimisation des diagrammes.

Des structures aléatoires peuvent être utilisées, mais les résultats corrects ne sont obtenus qu'après des tentatives multiples, sans qu'il soit possible d'assurer que la solution retenue soit la meilleure.

Les structures fractales donnent des résultats dont on sait déterminer, à priori, les caractéristiques. Certaines structures se prêtent particulièrement bien à la confection d'antennes et permettent d'ajuster au mieux la finesse du lobe principal et les remontées des lobes secondaires.

Des motifs fractals particuliers permettent de régler les problèmes des réseaux dans des espaces à 1, 2, ou 3 dimensions. Les structures fractales peuvent être obtenues par des processus déterministes mais aussi par l'introduction de fluctuations aléatoires autour du modèle déterministe, qui donnent des structures fractales pseudo-aléatoires ou des modèles construits

par des processus totalement aléatoires qui donnent des modèles fractals aléatoires.

### II. - GENERALITES -

#### II.1. Rappels sur les lois de rayonnement à l'infini.

Bien que la théorie développée ici soit applicable dans un espace de dimension 3, on se limitera dans cet exposé à un espace de dimension 2.

Si on considère une antenne dans l'espace à deux dimensions défini par les deux axes  $ox$ ,  $oy$  (figure 1) et une surface élémentaire  $dx dy$  pour laquelle la loi d'illumination est donnée par

$$I(x,y) = a(x,y) e^{j[\omega t + \Phi(x,y)]} \quad (1)$$

le champ produit en un point  $M$  situé à une distance éloignée de l'émetteur et repéré par ses coordonnées sphériques  $R$ ,  $\Phi$ ,  $\Theta$ , est donné par :

$$E(\Theta, \Phi, r) = f(r) \iint a(x,y) e^{-j[2\pi(ux+vy) + \Phi(x,y)]} dx dy \quad (2)$$

$$\text{où } u = \frac{\sin \Theta \cos \Phi}{\lambda}$$

$$\text{et } v = \frac{\sin \Theta \sin \Phi}{\lambda}$$

sont les fréquences spatiales.

On impose un rayonnement maximum dans une direction  $\Theta_0$ ,  $\Phi_0$ , par le choix de  $\Phi(x,y)$  et le champ rayonné s'exprime par

$$E(\Theta, \Phi, r) = f(r) \iint a(x,y) e^{-j2\pi(\Delta u x + \Delta v y)} dx dy \quad (3)$$

$$\text{où } \Delta u = u - u_0 \quad \text{et} \quad \Delta v = v - v_0$$

avec  $u_0 = \frac{\sin\theta_0 \cos\phi_0}{\lambda}$   $v_0 = \frac{\sin\theta_0 \sin\phi_0}{\lambda}$

Ce résultat connu exprime que le champ lointain est lié à la transformée de Fourier de la loi d'illumination.

On définit le gain d'antenne  $G$  par le rapport - généralement exprimé en dB - de la densité de puissance rayonnée dans un angle solide élémentaire  $d\Omega$  à celle qui serait rayonnée par une antenne isotrope qui produirait en  $M$  le champ  $E_0(\theta, \phi) = E_0$

$$G(\theta, \phi) = \frac{|E(\theta, \phi)|^2}{|E_0|^2} \quad (4)$$

Si  $P_0$  est la puissance rayonnée par chaque antenne

$$P_0 = \int_{4\pi} \frac{|E(\theta, \phi)|^2}{Z_0} r^2 d\Omega = \int_{4\pi} \frac{|E_0|^2}{Z_0} r^2 d\Omega$$

où  $Z_0$  est l'impédance d'onde du milieu, et alors

$$\int_{4\pi} G(\theta, \phi) d\Omega = 4\pi \quad (5)$$

relation qui montre que le volume limité par la surface du gain dans l'espace  $\Omega$  est constante : la réduction du gain en un lieu entraînera son augmentation ailleurs.

## II.2. Antennes en réseaux réguliers.

Si on considère les réseaux d'antennes, la loi d'illumination est discontinue et l'échantillonnage doit respecter la règle de Shannon pour éviter l'apparition des lobes d'ambiguïté.

Dans le cas de réseaux de surface réguliers constitués de  $M.N$  antennes disposées selon le schéma de la figure 2, le gain est donné par la relation connue

$$G(\theta, \phi) = g(\theta, \phi) (MN)^2 \frac{\sin^2 \pi N \Delta x \Delta u}{(N \sin \pi \Delta x \Delta u)^2} \frac{\sin^2 \pi M \Delta y \Delta v}{(M \sin \pi \Delta y \Delta v)^2} \quad (6)$$

où  $g(\theta, \phi)$  est le gain propre de chaque aérien.

La figure 3 montre les résultats obtenus pour un réseau linéaire constitué de 25 antennes ( $N=1$ ,  $M=25$ ) dans le cas où les distances entre antennes sont convenablement choisies et dans le cas où elles ne le sont pas.

Il doit être remarqué que :

- La largeur du lobe principal à 3dB est inversement proportionnelle à la dimension du réseau  $M \Delta y$ .

- L'affaiblissement apporté sur les lobes secondaires varie au voisinage du lobe principal en  $(1/\Delta v)^2$  pour un réseau linéaire (et en  $(1/\Delta v \Delta u)^2$  pour un réseau de surface), si bien que les lobes secondaires deviennent inutilement faibles.

- Le lobe d'ambiguïté qui apparaît résulte de la structure périodique du réseau qui rétablit la cohérence du rayonnement dans cette direction pour laquelle  $Dy \Delta v = 0 \bmod 2\pi$ .

Les structures régulières présentent donc, pour un nombre d'antennes fixé, deux inconvénients majeurs :

- La forme du diagramme de rayonnement imposée n'est pas idéale.

- Les lobes d'ambiguïté limitent l'espacement des antennes et, par suite, la finesse du lobe principal.

## II.3 Réseaux d'antennes à structure fractale.

La notion de géométrie fractale a été développée par Mandelbrot dans les années 1970, à la suite de travaux fondamentaux de nombreux mathématiciens, parmi lesquels on peut citer L.F. Richardson, G. Cantor, G. Peano, F. Hausdorff, A.S. Besicovitch, H. Minkowski, Bouligand.

Cette notion, qui réintroduit la rectifiabilité des objets non rectifiables par la mesure de Lebesgue, conduit à définir, pour un objet représenté dans un espace de dimension entière  $n$ , une dimension fractale réelle  $D$  inférieure à  $n$ .

La mesure est effectuée alors, comme pour la mesure de Lebesgue, par recouvrements de l'objet à l'aide de boules, mais leurs normes sont fractionnaires. Les boules sont de diamètres différents dans la mesure introduite par Hausdorff-Besicovitch, de diamètre identique dans celle introduite par Minkowski-Bouligand. La plupart du temps ces deux mesures coïncident.

Si on introduit une géométrie fractale dans le motif rayonnant d'un réseau d'antennes, la loi de rayonnement est toujours fournie par la relation (1) qui fait apparaître la transformée de Fourier de la loi d'illumination.

Il est facile de montrer que le rayonnement varie, en moyenne, selon la loi

$$G(\Delta u, \Delta v) = \frac{1}{(\Delta u, \Delta v)^D} \quad (7)$$

Bien qu'il s'agisse d'une loi statistique valable uniquement sous certaines conditions, elle montre que le choix de la dimension fractale fixe le niveau moyen des lobes secondaires, donc offre la possibilité de modifier la forme du diagramme de rayonnement.

Si l'antenne est construite dans un espace de dimension  $n$ , la finesse du lobe principal et le niveau des lobes secondaires seront d'autant plus importants que l'écart  $\Delta = n - D$  sera grand.

Pour les réseaux réguliers  $\Delta = 0$ . L'intérêt des structures fractales apparaît par le choix que l'on peut faire de  $\Delta$  qui devient alors un paramètre supplémentaire.

Cependant, certaines structures fractales peuvent laisser apparaître des lobes d'ambiguïté. Le choix des modèles est donc important. Des deux classes fractales habituellement considérées, auto-affines ou auto-similaires, seules les dernières sont considérées dans cette présentation, car elles apportent des solutions plus avantageuses.

L'élimination des lobes d'ambiguïté peut être obtenue par des fluctuations aléatoires à l'intérieur des structures déterministes. Elles peuvent être appliquées aux structures fractales. On obtient alors des structures fractales pseudo-aléatoires.

Une dernière méthode de génération de structure fractale est possible par un processus totalement aléatoire. Ces structures seront appelées fractales aléatoires. Cependant, des lois générales peuvent être désirées pour obtenir des propriétés statistiques des diagrammes de rayonnement. On peut alors utiliser une structure aléatoire pondérée par une loi de distribution déterministe de la géométrie désirée.

### III. - LES RESEAUX D'ANTENNES FRACTALS AUTO-SIMILAIRES -

Les modèles fractals auto-similaires sont caractérisés par le fait que toutes les parties de l'ensemble sont semblables à lui-même (similitude).

Ces modèles offrent des propriétés intéressantes dans les problèmes de rayonnement. On a cherché à utiliser ces propriétés pour la conception de réseaux lacunaires.

Après une présentation simple de tels ensembles, une relation générale est proposée et des exemples de

modèles montrent, par comparaison avec les réseaux classiques, l'avantage qu'on peut tirer de ces conceptions.

#### III.1. Un exemple simple : Le modèle triadique de Cantor.

Les ensembles triadiques de Cantor, dont la figure 4 rappelle le motif à différentes étapes, sont bien connus. On peut remarquer qu'il est possible de les représenter par l'union des intervalles définis par la relation :

$$C_n = U_j \left[ \frac{p_j, p_j + 1}{3^n} \right] \quad (8)$$

où  $U_j$  représente l'union des intervalles et

$$p_j = \sum_{i=0}^{n-1} a_{ij} 3^i$$

pour toutes les combinaisons possibles des  $a_{ij} \in \{0, 2\}$ .

Cet exemple bien connu de motif fractal auto-similaire a pour dimension :

$$D_{c3} = \frac{\ln 2}{\ln 3} = 0,631$$

#### III.2. Généralisation du modèle de Cantor.

Il est possible de généraliser la présentation de Cantor en définissant les ensembles

$$K_{1n} = U_j \left[ \frac{p_j, p_j + 1}{B^n} \right] \quad (9)$$

où  $B$  est réel et

$$p_j = \sum_{i=0}^{n-1} a_{ij} B^i$$

avec  $a_{ij} \in \{0, a_1, a_2 \dots a_i \dots a_{l-1}\}$  (10)

où  $a_i$  est réel inférieur à  $B - 1$  ( $a_i$  et  $B$  ne sont plus nécessairement des entiers),  $1 \leq B$ .

L'objet a alors pour dimension fractale :

$$D_K = \frac{\ln l}{\ln B} \quad (11)$$

ce qui permet d'ajuster la dimension aux valeurs souhaitées.

La figure 5 donne un exemple de motif pour  $B = 11$  et

$$a_{ij} \in \{0, 3, 5, 7, 10\}.$$

Enfin, on peut généraliser plus complètement encore la représentation en définissant le motif fractal par :

$$K_{2n} = U_j \left[ \frac{p_j \cdot p_j + c_{(aoj)}}{B^n} \right]$$

On obtient alors un motif non symétrique dont la figure 6 donne un exemple. Excepté le cas où  $C_{(aoj)}$  sont irrationnels, il est simple de voir qu'on peut obtenir tous les motifs en prenant  $C_{(aoj)} = 1$  par un choix convenable de  $B$  et  $I$ .

Tous ces motifs sont auto-similaires et possèdent des dimensions fractales inférieures à 1.

Si on définit le motif  $K_n$  déduit de  $K_{2n}$  par l'affinité  $B^n$  :

$$K_n = U_j [p_j, p_j + 1] \quad (12)$$

on décrit alors un motif dont l'étendue est  $B^n$  et composée de  $I^n$  segments.

La densité des segments, de longueur unitaire, à la dimension totale du motif est alors

$$d = \left( \frac{I}{B} \right)^n \quad (13)$$

Si on assimile les segments à des aériens,  $1/d$  représente la distance moyenne entre 2 aériens. Le réseau devient donc d'autant plus lacunaire que  $n$  est élevé.

### III.3. Réseaux d'antennes fractals auto-similaires.

Les réseaux d'antennes fractals auto-similaires présentent des caractéristiques de rayonnement intéressantes.

On supposera que le nombre d'antennes du réseau  $N$  est fixé a priori.

Les lois générales du rayonnement sont établies et il est analysé les niveaux des lobes secondaires et la largeur du lobe principal dont les relations avec la dimension fractale  $D$  sont étudiées.

#### III.3.1. Loi générale du rayonnement.

Pour les systèmes auto-similaires, il est plus intéressant d'exprimer cette loi en remarquant, à partir du modèle  $K_n$  que la loi d'illumination à l'étape  $n$ ,  $A_n(x, y) = a_n(x, y) e^{j\Phi(x, y)}$  est obtenue à partir de la loi  $A_{n-1}(x, y)$  par une auto-similarité définie par (figure 7) :

$$A_n(x, y) = \sum_{i=1}^I A_{n-1}(x, y) * \delta(x - x_{in}) * \delta(y - y_{in}) \quad (14)$$

où  $\delta(x)$  est la fonction de Dirac,  $*$  dénote le produit de convolution et qui donne le champ rayonné

$$E_n(\theta, \phi) = \sum_{i=1}^I E_{n-1}(\theta, \phi) e^{j2\pi \Delta u x_{in}} e^{j2\pi \Delta v y_{in}} \quad (15)$$

et si  $E_0(\theta, \phi)$  est le champ rayonné à l'étape 0

$$E_n(\theta, \phi) = E_0(\theta, \phi) \quad (16)$$

$$E_n(\theta, \phi) = E_0(\theta, \phi)$$

$$\prod_{k=1}^n \left[ \sum_i e^{j2\pi(\Delta u x_{ki} + \Delta v y_{ki})} \right]$$

Le terme

$$T_n(\theta, \phi) = R_n(\Delta u, \Delta v) =$$

$$\prod_{k=1}^n \left[ \sum_i e^{j2\pi(\Delta u x_{ki} + \Delta v y_{ki})} \right]$$

est le facteur de réseau de l'antenne dans lequel

$$S = \frac{x_{ki}}{x_{(k-1)i}} = \frac{y_{ki}}{y_{(k-1)i}}$$

est le rapport de similitude, égal, dans le cas du modèle  $K_n$  à  $B$ .

Le facteur de réseau s'exprime alors par la relation :

$$T_n(\theta, \phi) = R_n(\Delta u, \Delta v) =$$

$$\prod_{k=1}^n \left[ \sum_{i=1}^I e^{j2\pi S^k (\Delta u x_{0i} + \Delta v y_{0i})} \right]$$

ce qui donne

$$|T_n(\theta, \phi)|^2 = |R_n(\Delta u, \Delta v)|^2 = \quad (18)$$

$$\prod_{k=1}^n \left[ 1 + \sum_{i=1}^1 \sum_{h \neq i=1}^1 \cos 2\pi S^k [\Delta u(x_{ih}) + \Delta v(y_{ih})] \right]$$

$$\text{avec } x_{ih} = x_{oi} - x_{oh} \\ y_{ih} = y_{oi} - y_{oh}$$

Le facteur de réseau s'exprime de façon itérative par la relation

$$\frac{|R_n(\Delta u, \Delta v)|^2}{|R_{n-1}(\Delta u, \Delta v)|^2} = \quad (19)$$

$$1 + \sum_{i=1}^1 \sum_{h \neq i=1}^1 \cos [2\pi S^n [\Delta u(x_{ih}) + \Delta v(y_{ih})]]$$

qui montre que la relation est multiplicative et que le rapport est une fonction de similitude appliquée sur les arcs.

La figure 8 donne, pour l'exemple de la figure 5, les fonction  $|R_n(\Delta u, \Delta v)|^2$  qui s'expriment par la relation

$$R_n(\Delta u) = \prod_{k=1}^n [4 \cos(7B^k \pi \Delta u) \cos(3B^k \pi \Delta u) + 1]^2$$

Ces figures font ressortir l'aspect fractal du rayonnement exprimé en fonction de  $\Delta u$  ainsi que la réduction des lobes secondaires en fonction de  $\Delta u$  et de l'ordre  $n$  du motif.

### III.3.2. Protection sur les lobes secondaires.

Dans le cas de la structure  $K_n$  l'expression (18) montre que :

- Dans la direction du lobe principal

$$|T_n(\theta_0, \phi_0)|^2 = |R_n(0, 0)|^2 = 1^{2n}$$

- Dans les autres directions

$$|T_n(\theta_0, \phi_0)|^2 = |R_n(\Delta u, \Delta v)|^2 = 1^n$$

La protection sur les lobes secondaires, dans le cas du modèle  $K_n$ , est donc donnée par

$$\alpha = 1 - n = B \cdot nD \quad (20)$$

On retrouve donc la dépendance de la protection en fonction de la dimension fractale.

### III.3.3. Largeur du lobe principal d'antenne.

La largeur du lobe principal d'un réseau d'antennes, dans une direction donnée, est inversement proportionnelle à l'étendue du réseau dans cette direction.

Le nombre d'antennes du réseau est égal à  $N = I^n$  et la distance moyenne, pour un réseau linéaire de type  $K_n$  est  $\Delta_0/d$  où  $\Delta_0$  est la distance minimale entre deux antennes.

La dimension du réseau vaut alors

$$L_F = \frac{N \Delta_0}{d} = \Delta_0 B^n$$

Pour un réseau régulier, la dimension vaut :

$$L_R = \Delta_0 \cdot N = \Delta_0 I^n$$

Le rapport des largeurs des lobes principaux du réseau fractal  $B_F$  au réseau régulier  $B_R$  est donc

$$\frac{B_F}{B_R} = \beta \frac{L_R}{L_F} = \beta D^n$$

où  $\beta$  est un coefficient supérieur mais voisin de 1, prenant en compte l'effet lacunaire du réseau fractal.

La figure 9, qui représente les variations des paramètres définissant le lobe principal et les lobes secondaires, montre que le choix de la dimension fractale ne peut résulter que d'un compromis entre la finesse du lobe principal et les niveaux des lobes secondaires.

En effet, pour un ordre  $n$  donné, il apparaît :

- Que le rapport  $B_F/B_R$  est une fonction décroissante de  $D$ .
- Que la protection  $\alpha$  sur les lobes secondaires est une fonction croissante de  $D$ .

L'utilisateur, en fonction de ses applications, doit donc choisir la protection sur les lobes secondaires et la finesse du lobe principal. Il en déduit alors la dimension fractale et l'ordre du motif.

### IV. - EXEMPLE DE RESEAUX -

La réalisation de réseaux d'antennes fractals peut être entreprise avec les formes de réseaux les plus diverses.

On présente dans cette partie quelques exemples permettant d'illustrer les développements précédents en utilisant, dans les figures 10, 11, 12 un réseau

linéaire rectiligne pour permettre une comparaison simple avec un réseau régulier.

Quatre réseaux ont été utilisés pour cet exemple (figure 10). Les données communes, nécessaires pour la comparaison, sont :

- Le nombre d'aériens par réseau N, égal selon les cas à 5, 25, 125 ou 625.
- La direction du lobe principal  $\Phi_0 = 90$ ,  $\Theta_0 = 180$  degrés (qui donne évidemment un lobe arrière à 0 degré).
- La fréquence d'émission  $F = 15\text{MHz}$ , pour laquelle  $\lambda = 20\text{m}$ .
- La distance minimale entre antenne est  $D_{\text{min}} = \lambda/2$ .

#### Réseau A

Réseau linéaire rectiligne composé d'antennes espacées de  $\lambda/2 = 10\text{ m}$ , correspondant à un espacement  $D_y$  maximum pour éviter des lobes d'ambiguïté.

Pour ce réseau

$$B = 5$$

$$a_i \in \{0 - 1 - 2 - 3 - 4\}$$

$$\text{Dimension fractale} \quad D = 1.$$

#### Réseau B

Réseau linéaire fractal avec :

$$B = 5,7$$

$$a_i \in \{0 - 1,37 - 2,41 - 3,43 - 4,59\}$$

$$\text{Dimension fractale} \quad D = 0,877.$$

#### Réseau C

Réseau linéaire fractal avec :

$$B = 6,27$$

$$a_i \in \{0 - 1,47 - 2,65 - 3,77 - 5,37\}$$

$$\text{Dimension fractale} \quad D = 0,797.$$

#### Réseau D

Réseau linéaire fractal avec :

$$B = 7,87$$

$$a_i \in \{0 - 2,07 - 3,53 - 4,86 - 6,83\}$$

$$\text{Dimension fractale} \quad D = 0,635.$$

La figure 11 montre l'influence de la dimension fractale sur les lobes secondaires. On doit noter que du réseau régulier pour lequel  $D = 1$  au réseau fractal de dimension minimale  $D = 0,635$

- Les remontées des lobes secondaires augmentent.
- La largeur du lobe principal diminue.

La figure 12 montre l'influence de la dimension fractale sur la largeur du lobe principal. Les résultats présentés ici sont les parties des diagrammes de rayonnement de la figure 10 représentant les angles de rayonnement de 160 à 200 degrés. Pour un même nombre d'antennes on constate que la largeur du lobe principal diminue avec la dimension fractale. On a reporté les valeurs de la largeur théorique de  $B_F$  calculée à partir de  $B_R$  par la relation

$$B_F = B_R D^n$$

On peut constater que cette relation est très bien vérifiée.

La figure 13 enfin, met en évidence l'effet de l'ordre du motif  $n$  qui varie de 1 à 4. Le nombre d'antennes prend donc les valeurs respectives  $N = 5, 25, 125, 625$ .

On constate le bon accord entre la théorie et les résultats puisque l'on remarque :

- Une diminution des lobes secondaires lorsque  $n$  augmente.
- Une réduction de la largeur du lobe principal lorsque  $n$  augmente.

Les réseaux à deux dimensions traités dans le cas général présentent évidemment des propriétés semblables. La figure 14 représente des diagrammes de rayonnement pour des réseaux fractals composés de 125 antennes dans lesquels la disposition des antennes est à l'intérieur d'un cercle.

La figure 14a représente le diagramme de rayonnement d'un réseau linéaire circulaire régulier tel que la distance entre deux antennes soit  $\lambda/2$ . La loi d'illumination connue montre une remontée des lobes secondaires importante.

Les figures 14b et 14c montrent les diagrammes obtenus de deux réseaux fractals dont les dimensions sont voisines de  $D = 1,6$ . Ils révèlent l'avantage que l'on peut trouver dans les réseaux de surface à structure fractale, tant sur la réduction du lobe principal que sur la protection des lobes secondaires.

La figure 15 donne les diagrammes de rayonnement de réseaux de surface fractals inscrits dans des rectangles. Ces réseaux, dont les dimensions sont voisines de  $D = 1,6$ , ont été obtenus en combinant une loi fractale et une loi aléatoire. On peut constater encore la bonne qualité des diagrammes de rayonnement obtenus avec 125 antennes, mais l'introduction de la loi aléatoire ne permet pas de contrôler, à priori, les niveaux des lobes secondaires.

Dans le premier cas, la protection sur le lobe arrière est assurée, alors qu'elle ne l'est pas dans le cas du réseau régulier.

## V. - CONCLUSION -

La conception de réseaux d'antennes lacunaires de grande dimension a été abordée à travers une théorie fractale, par la construction de motifs autosimilaires étendus bien adaptés au problème.

La formulation proposée conduit à des résultats en bon accord avec la théorie, accord que confirment les exemples donnés.

Le compromis entre la largeur du lobe principal et la protection à assurer sur les lobes secondaires fixe la dimension fractale.

La méthode proposée consiste à établir le réseau à partir d'un (réseaux linéaires) ou deux (réseaux de surface) systèmes de numération lacunaires dont les bases et les chiffres sont fractionnaires.

Les motifs dépendent du choix que l'on peut alors faire et il convient de choisir les chiffres du système de numération de façon à ne pas introduire, contrainte facile à respecter, des régularités dans leur progression.

L'introduction des variations aléatoires, ou de motifs fractals totalement aléatoires, ou toute combinaison de ces méthodes, peut être utilisée. Les résultats constatés n'ont pas mis en évidence d'amélioration des diagrammes.

Les diagrammes obtenus peuvent, comme pour les réseaux réguliers, être améliorés en utilisant des lois de pondération sur l'amplitude complexe des signaux de chaque antenne.

La structure purement fractale paraît donc préférable puisqu'elle conserve un aspect totalement déterministe en organisant, de façon rigoureuse, le désordre.

---

## BIBLIOGRAPHIE

- [1] A. Besicowitch - Almost periodic functions.  
Cambridge U. press. 1932.
- [2] B.B. Mandelbrot - Fractals, form and  
dimension.  
Sans Francisco. 1977.

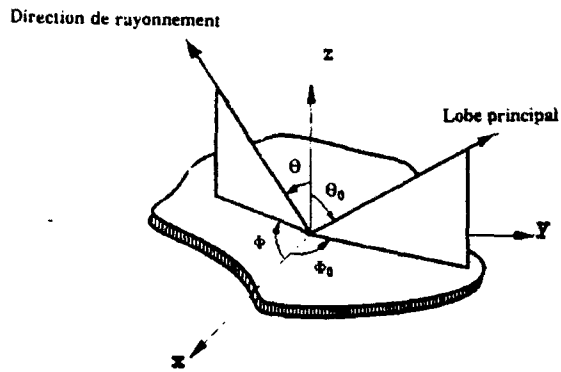


FIGURE 1 : Géométrie du rayonnement d'une surface définie par sa loi d'illumination.

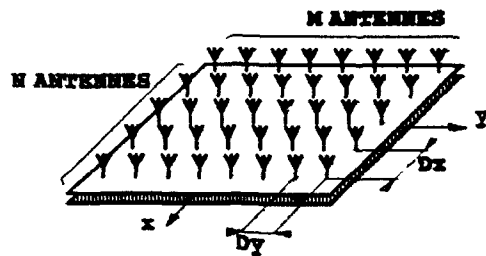


FIGURE 2 : Réseau de surface régulier.

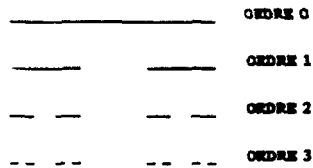


FIGURE 4 : Ensemble triadique de Cantor.

$$K_{1n} = U_j \left[ \frac{p_j \cdot p_{j+1}}{B^n} \right]$$

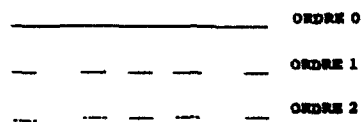


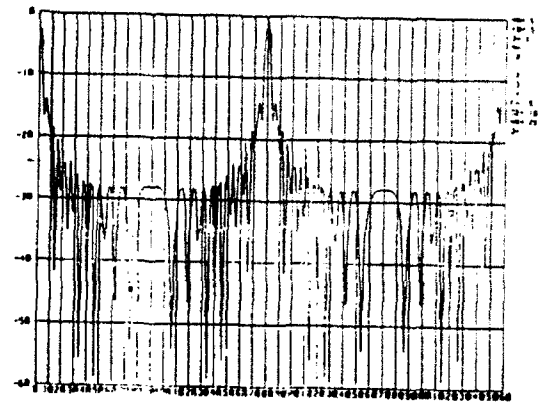
FIGURE 5 : Motifs fractals définis par

$$B = 11, \quad I = 5$$

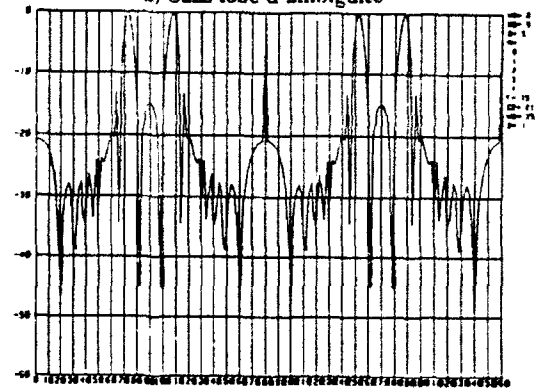
$$a_i \in \{0, 3, 5, 7, 10\}$$

$$D = 0,454$$

FIGURE 3 : Exemple de diagramme d'un réseau linéaire régulier



a) Sans lobe d'ambiguïté



b) Avec lobe d'ambiguïté

$$K_{2n} = U_j \left[ \frac{p_j \cdot p_j + c(a_0)}{B^n} \right]$$



FIGURE 6 : Motifs fractals définis par

$$B = 11, \quad I = 5$$

$$a_i \in \{0, 3, 5, 7, 11\}$$

$$C_{a_i} \in \{1, 0,5, 1, 0,5, 1\}$$



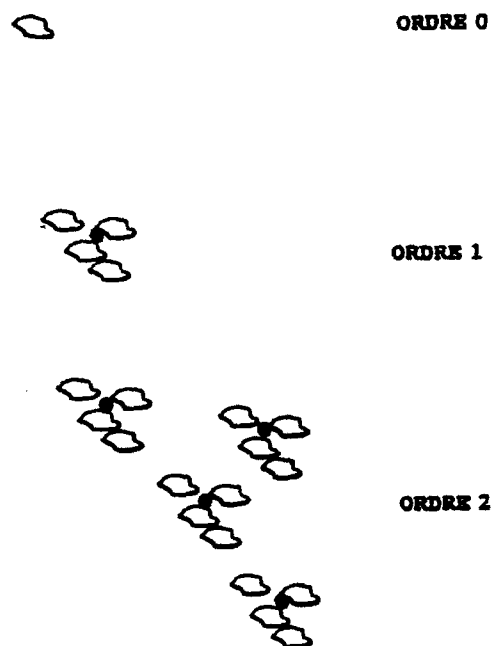


FIGURE 7 : Configuration générale d'une antenne fractale autosimilaire dans un espace à 2 dimensions.

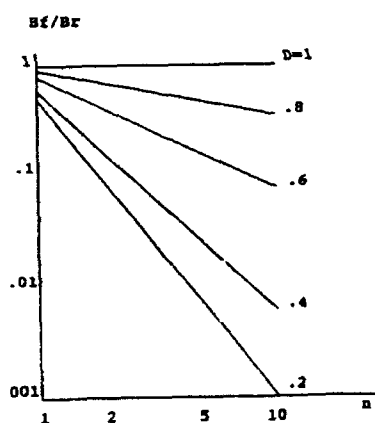


FIGURE 9a : Variation du rapport  $B_f/B_r$  en fonction de l'ordre du motif

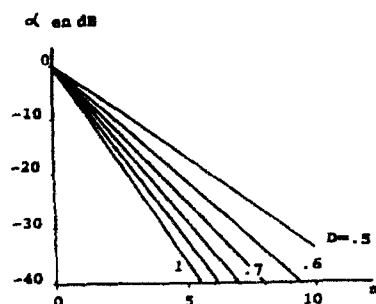


FIGURE 9b : Variation de la protection sur les lobes secondaires en fonction de l'ordre du motif pour  $B = 5$

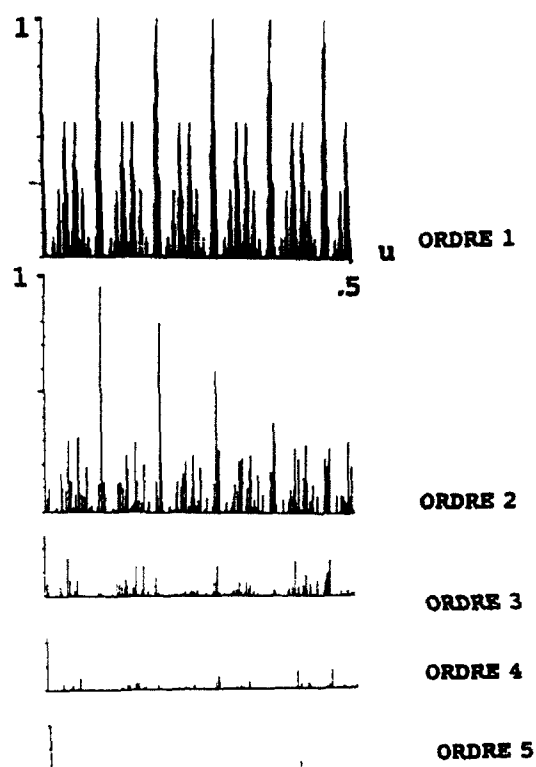


FIGURE 8  
VARIATION DU FACTEUR DE RESEAU

$$|R(\Delta u)|^2$$

$$B=11 \quad a \in \{0, 3, 5, 7, 10\}$$

RESEAU A    •   •   •   •   •

DIMENSION  $D=1$

RESEAU B    •   •   •   •   •

DIMENSION  $D=.877$

RESEAU C    •   •   •   •   •

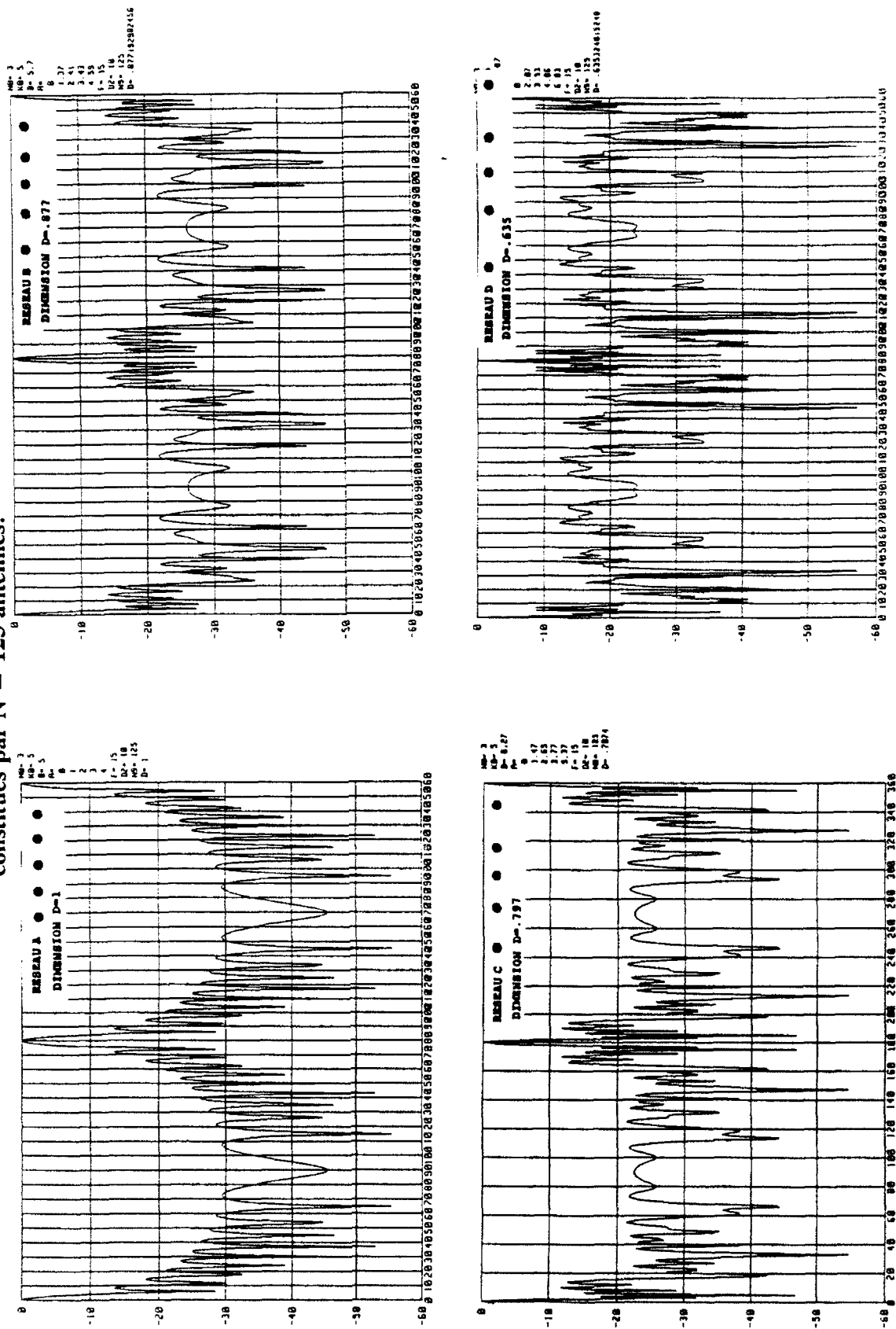
DIMENSION  $D=.797$

RESEAU D    •   •   •   •   •

DIMENSION  $D=.635$

FIGURE 10 : Motifs fractals de base (ordre 1) des réseaux linéaires utilisés en exemple.

FIGURE 11 : Influence de la dimension fractale sur les lobes secondaires pour des réseaux constitués par  $N = 125$  antennes.



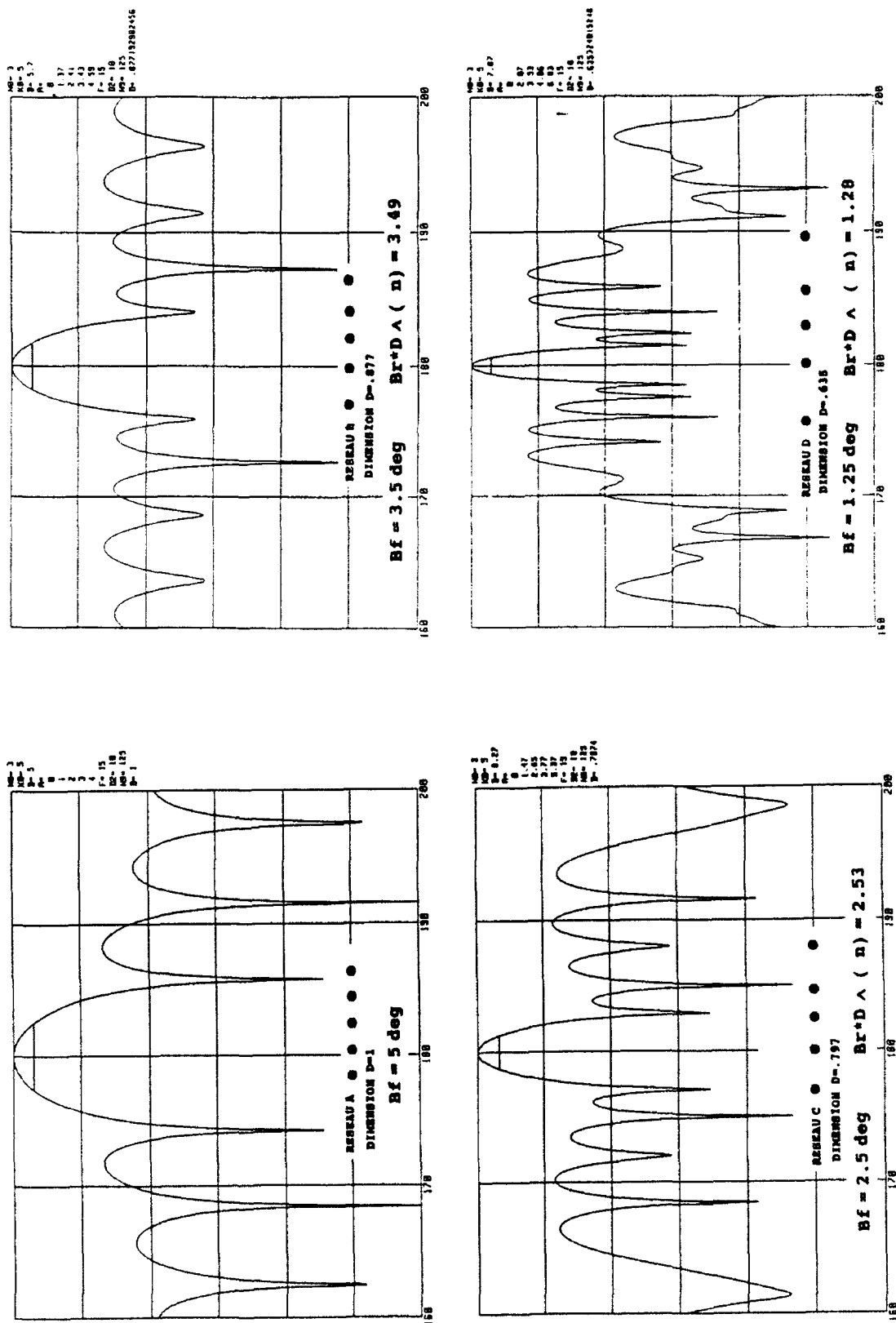
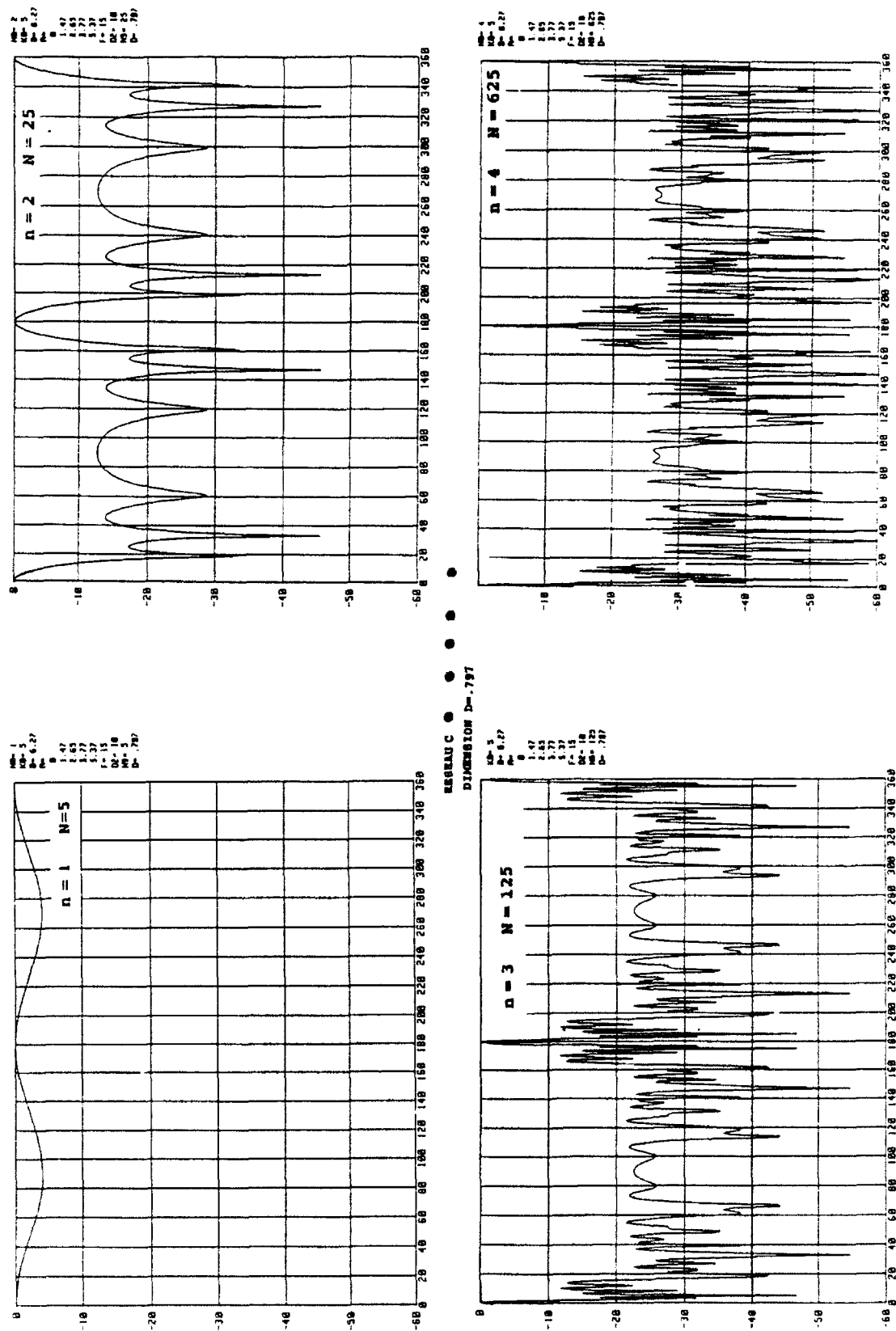


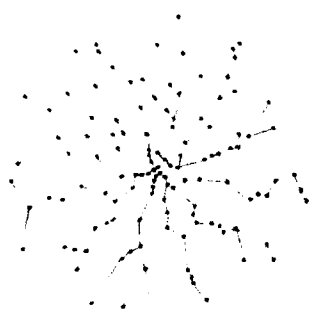
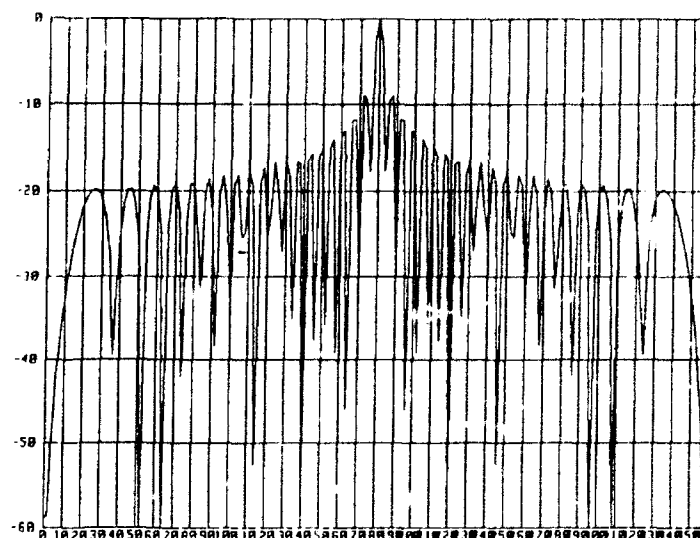
FIGURE 12 : Influence de la dimension fractale D sur la largeur du lobe principal pour des réseaux constitués par 125 antennes.



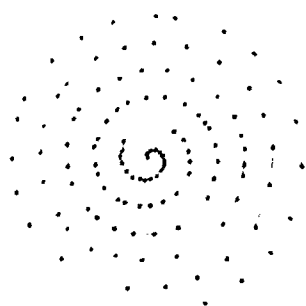
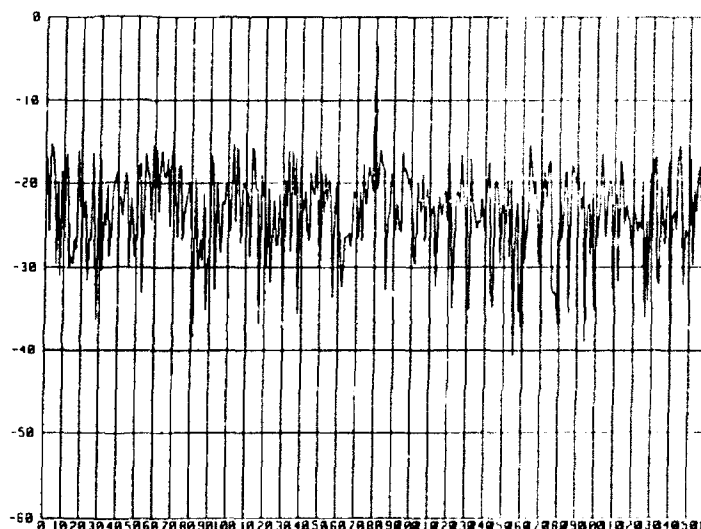
**FIGURE 13 : Influence de l'ordre  $n$  du motif sur le diagramme de rayonnement.**  
Réseau C de dimension  $D = 0,797$



(a)



(b)



(c)

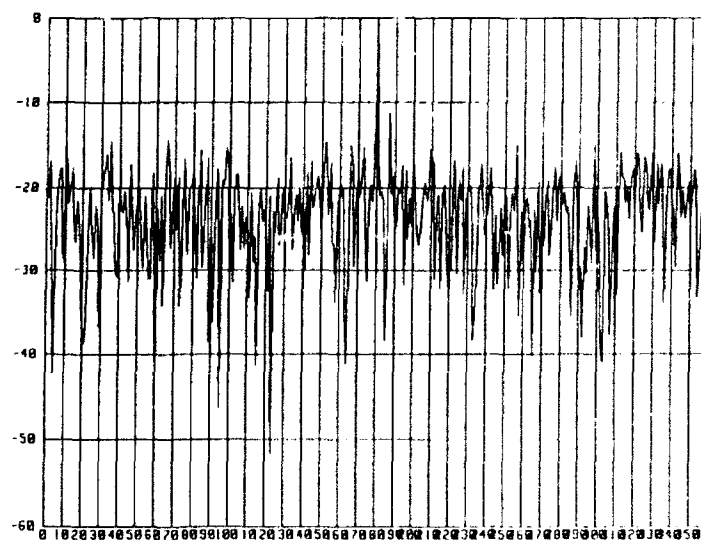
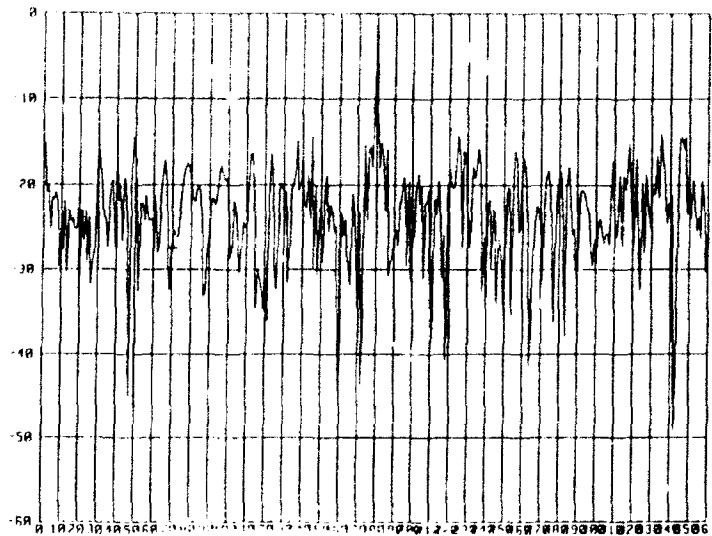


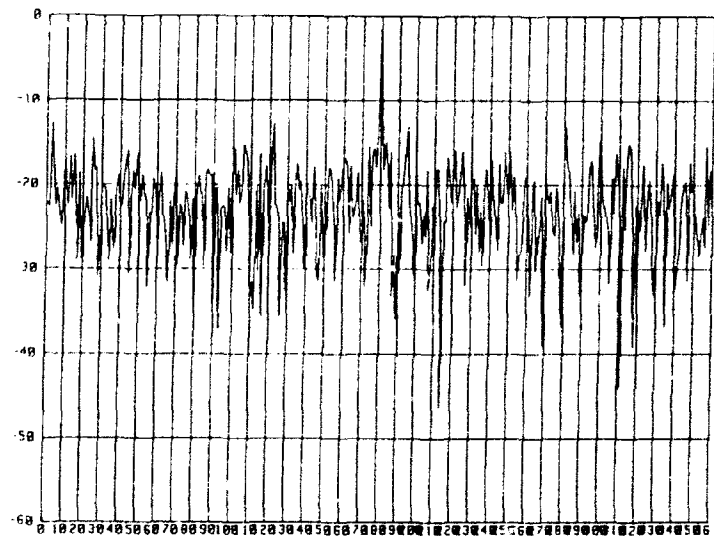
FIGURE 14 : Réseaux fractals "circulaires" à  
 $N = 125$  antennes.  
 a) Réguliers b) c) Fractals



(a)



(b)



(c)

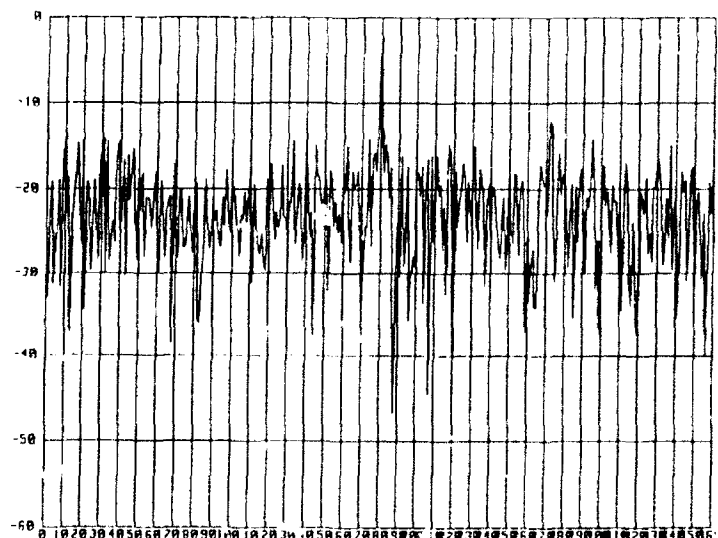


FIGURE 15 : Réseaux fractals "rectangulaires" à  
 $N = 125$  antennes.

## DISCUSSION

### L. LEWANDOWSKI

In the case of random arrays, one can use Ruze's statistical method of estimating probability of side-lobe-level exceedance. Is there a comparable way to "guesstimate" the side-lobe-level peak exceedance for arrays determined using the fractal design approach?

### AUTHOR'S REPLY

La différence entre l'approche purement aléatoire et l'approche fractale vient du fait que la détermination fractale est déterministe et que par conséquent le diagramme de rayonnement peut être exprimé de façon analytique. On peut donc parfaitement contrôler les lobes secondaires (et tout le diagramme de rayonnement) du réseau. Dans les réseaux aléatoires on ne peut avoir qu'une approche statistique, ce qui ne permet pas de garantir les remontées locales.

*The difference between the purely random approach and the fractal approach is explained by the fact that fractal determination is deterministic and that as a result the antenna pattern can be expressed in analytical fashion. This means that the side lobes (and the whole of the antenna pattern) of the array can be completely checked. With random arrays we can only adopt a statistical approach which does not enable us to allow for local variations.*

### C. BOND

Durant votre exposé, vous avez fait référence à un grand nombre d'éléments à l'intérieur de votre antenne. Croyez-vous que cette méthode peut être utile à la bande HF où l'on utilise, en général, de sept à onze antennes (éléments)?

*During your talk, you referred to a large number of elements inside your antenna. Do you think that this method could be useful for the HF band, where some 7 to 11 antennas (elements) are commonly used?*

### AUTHOR'S REPLY

Cette méthode a été étudiée pour des réseaux de grande dimension. Elle peut être appliquée approximativement pour des réseaux dépassant 25 antennes. En HF, elle peut être utilisée pour ces réseaux sachant que l'on utilise des réseaux de grande dimension, notamment en radar transhorizon.

*This method was designed for large scale arrays. It can be applied approximately to arrays with over 25 antennas. It can be used for HF arrays, given that large scale HF arrays are used, in particular for over-the-horizon radar.*

## LOCALISATION D'EMISSIONS A EVASION DE FREQUENCE ET RAFALES PAR CAPTEUR A ANALYSE SEQUENTIELLE EN BALAYAGE RAPIDE

D. JOSSET  
THOMSON-CSF Division RGS  
66, rue du Fossé Blanc  
92 231 GENNEVILLIERS  
FRANCE

### Résumé :

L'utilisation de techniques de transmission par évocation de fréquence ou rafales en VUHF afin de lutter contre les systèmes ECM classiques se répand de plus en plus. L'apparition de capteur goniométriques rapides rend possible l'interception et la goniométrie de ces signaux.

La localisation de ces signaux par un système classique nécessite une mise en attente sur l'émission. Afin d'éviter ce problème, on propose de disposer un système d'extraction des émissions à évocation de fréquence et rafales avec les goniomètres avant la localisation, celle-ci s'effectuant sur des émissions "synthétiques" calculées à partir des détections goniométriques.

Le calcul de ces émissions "synthétiques" nécessite un algorithme permettant la reconnaissance et l'identification des types d'émissions. A partir des reconnaissances effectuées sur plusieurs goniomètres, on recherche les émetteurs ayant les mêmes caractéristiques sur les différents capteurs avant d'en effectuer la localisation par une méthode des moindres carrés cartésiens.

### 1 - MODELISATION DE L'ENVIRONNEMENT

Afin de mesurer les performances d'un système de localisation d'émissions à évocation de fréquence et rafales à partir de capteur à analyse séquentielle, il a été nécessaire de développer une simulation de l'environnement radioélectrique ainsi que du goniomètre.

a) intercepteur goniométrique

Les nouvelles générations d'intercepteur goniométrique permettent l'interception des émissions à évocation de fréquence. Ces intercepteurs possèdent des vitesses de balayage dépassant 1 GHz par seconde. Ces intercepteurs utilisent le principe de détection d'énergie afin de n'effectuer une mesure goniométrique que lorsque le canal radioélectrique est actif.

L'intercepteur modélisé dans le cadre de cette étude est le TRC 612 de THOMSON-CSF. Les principes du système de localisation proposé peuvent s'appliquer à tous les types de goniomètre utilisant les mêmes principes de mesure (mesure goniométrique sur activité d'un canal).

Le fonctionnement de ces intercepteurs sont de par leur conception asynchrones; le temps de scrutation d'une bande dépend fortement de la densité de l'environnement.

b) environnement

L'environnement cible est un environnement champs de bataille. Sur une zone donnée (50 km \* 50 km), on dispose d'un ensemble de 100 réseaux de 10 émetteurs. La dotation en émetteur à évocation de fréquence est variable. De même pour les émetteurs du type rafales. La répartition géographique de ces émetteurs suit une loi aléatoire uniforme à 2 dimensions dans la zone considérée.

c) communications

Les modèles de communications pris en compte pour les communications en mode fréquence fixe ( FF ) et en mode évocation de fréquence ( EVF ) sont des communications parlées. La modélisation des alternats est décrite en annexe 2.

Les émissions EVF sont modélisées par un tirage aléatoire uniforme des paliers de fréquence



sur domaine fréquentiel déterminé. Dans l'application présentée ci-après, le domaine fréquentiel est réduit à une seule bande d'étalement. L'utilisation d'une distribution EVF répartie sur plusieurs sous-bandes d'étalement ne modifiera pas l'algorithme de localisation.

Les rafales seront assimilées à des transmissions de données ponctuelles. Elles émettent en mode fréquence fixe. La durée des émissions rafales est comprise entre 100 ms et 1 seconde avec répétition possible.

#### d) propagation

Le modèle de propagation pris en compte est un modèle de propagation en espace libre. L'objectif de cette étude étant la validation d'une méthode face à un environnement donné, les perturbations liées à la propagation par onde de sol ne seront pas pris en compte (effets de réduction de l'environnement réellement visible au niveau de chaque capteur).

## 2 - RECONNAISSANCE DES EMISSIONS FF / EVF / RAFALES

#### a) mode de fonctionnement du goniomètre

L'intercepteur goniométrique effectue un balayage sur le domaine fréquentiel déterminé [ $f_{\min}$ ,  $f_{\max}$ ]. A chaque détection d'énergie, il effectue une mesure goniométrique sur l'émission radioélectrique. La détection d'énergie dans un canal radioélectrique s'effectue dans un temps très bref (typique : 10  $\mu$ s). Par contre, la mesure goniométrique nécessite une présence du signal sur environ 500  $\mu$ s.

Pendant une durée d'observation T, le goniomètre aura effectué plusieurs balayages successifs de la bande [ $f_{\min}$ ,  $f_{\max}$ ] considérée.

#### b) discrimination des émissions

Chaque émission radioélectrique possède un vecteur de paramètres discriminants; ce vecteur est caractéristique de l'émission. Pour une émission FF, la caractérisation peut être représentée par le vecteur suivant:

( fréquence centrale, largeur d'émission, modulation [types, ...], azimut, instants d'émission )

Il en sera de même pour une émission EVF :

( bandes de fréquences, loi de distribution des paliers, modulation (type, ...), azimut, instants d'émission ).

En goniométrie séquentielle, on ne dispose pas de la mesure de l'ensemble de ces paramètres discriminants. La goniométrie limite le nombre de paramètres mesurés à 4 qui sont:

- fréquence
- azimut
- instant de mesure
- niveau.

Le critère niveau dépend fortement des conditions de propagation et des modulations; sa valeur n'aura qu'une valeur subalterne. Ainsi il ne sera pris en compte dans la discrimination que sur des écarts importants lorsque les autres paramètres seront semblables.

L'identification d'une émission FF se fera essentiellement sur le critère fréquentiel. Les critères d'azimut et de durée serviront à caractériser l'émission; ces paramètres seront utilisés en corrélation de localisation.

Pour une émission EVF, ce paramètre ne pourra plus être utilisé comme discriminant (notamment dans le cas des goniomètres séquentiels). Le paramètre discriminant essentiel sera la mesure azimutale de l'émission. A partir de ce critère discriminant, on déterminera les zones fréquentielles d'étalement des EVF (ce critère deviendra discriminant en corrélation de localisation).

#### c) reconnaissance des émissions

A partir des données acquises par le goniomètre, il existe de multiples méthodes de classification des données pour résoudre le problème posé. La méthode présentée ci-après est élaborée afin de tenir compte du caractère temps réel des opérateurs et limiter le coût en puissance de calcul.

La mesure goniométrique est constituée par le vecteur

( azimut,  
fréquence,  
date,  
niveau ).

Le classement des mesures sur une période d'extraction est effectuée suivant l'ordre suivant:

- 1) fréquence
- 2) azimut
- 3) date.

Le niveau ne permet que dans quelques occasions limitées de séparer 2 émissions.

Le nombre de mesures effectuées sur une émission dépend de la densité d'environnement et de la bande balayée. La durée de scrutation d'une bande étant très inférieure à la durée d'extraction, le goniomètre effectuera n mesures sur une même

émission FF. La construction d'un histogramme en fréquence permet par seuillage d'en extraire les émissions de type FF. La distinction des alternats de communications FF est établie par un histogramme en azimut; chaque alternat de communication est identifié par une datation au sein de la période d'extraction.

En effectuant un regroupement de toutes les mesures relatives à une même émission FF, on établit un plot "synthétique" représentatif de l'émission. Il est caractérisé par les paramètres suivants :

```
(
  type FF
  fréquence centrale
  azimut moyen
  écart_type en azimut
  date de début
  date de fin
)
```

Le discriminant fréquentiel ne pouvant être employé pour reconnaître les EVF, on effectuera une discrimination azimutale. Cette opération est effectuée après élimination des émissions FF. La détection des émissions EVF se fait par seuillage sur un histogramme en azimut des détections.

Sur un même azimut, plusieurs émissions EVF peuvent coexister (alignement dans les visées goniométriques). Aussi on identifie chaque émission EVF par la (ou les) zone(s) fréquentielle(s) occupée(s). Chaque plot EVF est alors caractérisé par un domaine d'étalement fréquentiel [f\_min, f\_max].

Les détections relatives à chaque émission EVF sont alors regroupées dans un plot "synthétique" représentatif de l'émission:

```
(
  type EVF,
  fréquence min,
  fréquence max,
  azimut moyen,
  écart-type azimutal,
  date de début,
  date de fin
)
```

La datation des émissions se fait en calculant la date inférieure et supérieure des détections relatives à un plot.

La reconnaissance des émissions de type rafale est faite sur le même principe que la reconnaissance FF. On établit la reconnaissance rafale sur l'ensemble des détections n'ayant pas servi aux reconnaissances précédentes. Cette opération n'est effectuée que sur les émissions rafales émettant en fréquence fixe.

On effectue l'identification de l'émission rafale au même titre que les émissions FF par un plot "synthétique" ayant les mêmes caractéristiques.

### 3 - PRINCIPES DE LOCALISATION

#### a) technique de localisation

La méthode de localisation utilisée est une méthode de localisation nécessitant l'emploi de plusieurs observateurs. La technique de localisation est une localisation par les moindres carrés cartésiens [1]. Nous rappelons brièvement les principes de la méthode en annexe 1.

#### b) techniques de corrélation entre capteur

Chaque capteur goniométrique effectue un ensemble de relevé de mesures sur les émissions radioélectriques interceptables. Chaque mesure est décrite par un vecteur associé

```
(
  capteur
  azimut
  fréquence
  niveau
  date
)
```

Dans les systèmes de localisation actuels, la corrélation entre capteur d'une émission est essentiellement basée sur le critère fréquentiel.

L'utilisation de ce critère comme discriminant dans les systèmes de localisation des émissions à évocation de fréquence nécessite une parfaite synchronisation des balayages des stations (interception des mêmes paliers au mêmes instants). Ce problème deviendra d'autant plus insoluble que les durées de paliers seront courtes.

Pour réduire l'influence du problème de synchronisation, on va effectuer avant la localisation une reconnaissance des émissions appelée extraction; chaque émission sera identifiée par un ensemble de paramètres

```
(
  capteur
  type (FF,EVF,RAF)
  azimut moyen
  sigma azimut
  fréquence centrale
  largeur de bande
  date de début
  date de fin
)
```

A partir des ensembles de plots représentatifs des émissions détectées, la première étape de la localisation est d'effectuer la corrélation des émissions entre les différents capteurs.

Les critères de corrélation seront classés dans un ordre décroissant :

```
(
  type
  fréquence
  date
  azimut
)
```

On dispose au niveau de la station de

localisation de n ensembles de mesures des émissions relatifs aux environnements perçus par chaque station de localisation. Nous appliquerons la corrélation entre émissions en respectant l'ordre de discrimination présentée ci-dessus.

Les émissions reconnues de même type au niveau extraction seront corrélées entre elles. Ensuite sera pris en compte le critère fréquentiel.

La corrélation des émissions de type FF est faite essentiellement sur la fréquence, ensuite sur les dates. La datation permettra de discriminer les alternats de communications FF. La nature asynchrone des goniomètres ne permet pas d'avoir les mêmes dates de début et fin des émissions entre goniomètres. Aussi la corrélation des dates d'émission entre les différents capteurs sera faite avec un degré d'incertitude. Ce degré d'incertitude prend en compte les effets des balayages non synchrones des différents goniomètres, ainsi que les erreurs de datation.

L'azimut n'est pas un critère à proprement dit de corrélation; il permet seulement de lever les ambiguïtés de localisation, en particulier lors de tirs goniométriques parallèles.

Pour les émissions de type EVF, on reprendra les mêmes discriminants. Etant donné que les caractéristiques des EVF mesurées au sein des divers goniomètres sont différentes (asynchronisme des balayages), on effectuera une corrélation en imposant une marge d'incertitude sur la corrélation. Le degré d'incertitude sur le domaine commun fréquentiel dépend des modes de fonctionnement (durée d'extraction, densité d'environnement,...). Ce degré d'incertitude est établi expérimentalement.

La corrélation sur les dates est effectuée avec la même méthode que celle employée en FF.

Les émissions rafales extraites sont analysées en respectant les principes décrits en fréquence fixe.

#### 4 - PERFORMANCES

Les performances de la localisation dépendent du temps d'observation du système et du nombre d'émetteurs actifs pendant cette période. On définit ainsi plusieurs paramètres de localisation, qui sont:

- . le taux de localisation représentant le nombre de localisations par cycle d'extraction sur le nombre d'émetteurs actifs,
- . le taux de localisation réussie représentant le nombre d'émetteurs correctement localisés.

On mesure une performance de

localisation par cycle d'extraction représentant le nombre de localisations correctes par rapport au nombre d'émetteurs actifs. La performance globale du système sera mesurée par le nombre d'émetteurs localisés sur N cycles d'extraction.

Avec 177 réseaux de 10 émetteurs ayant un taux d'activité de 50% et les modèles de communications décrits précédemment (80 % FF, 10% EVF et 10 % rafales) répartis sur la bande (30 MHz, 88 MHz), on obtient statistiquement par cycle d'extraction (cycle d'extraction de 2 secondes), des taux de localisation de :

- . 72 % pour les émetteurs FF actifs durant ce cycle  
soit 51 émetteurs sur 71 toutes les 2 secondes,
- . 33 % pour les émetteurs EVF actifs durant ce cycle  
soit 6 émetteurs sur 18 toutes les 2 secondes,
- . 41 % pour les émetteurs rafales actifs durant ce cycle  
soit 7 émetteurs sur 18 toutes les 2 secondes.

L'application d'un suivi de localisation sur plusieurs cycles d'extraction successifs permet de localiser tous les réseaux. Statistiquement sur une période de 1 minute (30 cycles d'extraction de 2 secondes), tous les émetteurs actifs auront été localisés.

Le système de localisation après extraction est conçu pour permettre un suivi temps réel des émissions. Aussi on pourra développer au sein de la station de localisation, un système de suivi temps réel des émetteurs; ceci permettra l'analyse temps réel de l'environnement radioélectrique.

#### 5 - CONCLUSION

Ce système de localisation a été développé pour s'adapter à l'intercepteur goniométrique TRC 612 de THOMSON-CSF afin de répondre aux nouvelles menaces EVF et rafales. La réalisation de ce système est en cours d'implémentation.

#### Bibliographie :

- [1] STATISTICAL THEORY OF D.F. FIXING. R.G. STANSFIELD - PEE 1947 Vol 94.
- [2] CLASSIFICATION AUTOMATIQUE DES DONNEES, Environnement statistique et informatique. CELEUX, DIDAY, GOVAERT, LECHEVALLIER, RALAMBONDRAINY - DUNOD.

## ANNEXE 1 : PRINCIPES DE LA LOCALISATION

Un goniomètre G de coordonnées (Xg, Yg) dans un repère (O, x, y) mesure l'angle sous lequel il intercepte une cible radioélectrique C de coordonnées (Xc, Yc). L'angle théorique (thétha) est donné par

$$\theta = \arctg \left( \frac{(Ys - Yg)}{(Xs - Xg)} \right)$$

La mesure est affectée d'un bruit inhérent au capteur employé. Nous prendrons comme hypothèse que l'erreur de mesure sur l'angle suit une loi normale de moyenne nulle. La probabilité d'effectuer une mesure avec une erreur phi autour theta s'écrit :

$$p(\Phi) d\Phi = \frac{1}{\sigma_\Phi \sqrt{2\pi}} \exp\left(-\frac{\Phi^2}{2\sigma_\Phi^2}\right) d\Phi$$

$\sigma_\Phi$  : écart-type de mesure du goniomètre  
Soit P la projection orthogonale de la cible sur le tir. On a

$$P_j = \overline{PC} = D_j g(\Phi)$$

avec D<sub>j</sub> la distance cible gonio

Pour N goniomètres, les erreurs de mesure d'angle sont indépendantes. En conséquence, la probabilité résultante sera le produit des probabilités élémentaires :

$$P(\Phi_1, \dots, \Phi_N) d\Phi_1 \dots d\Phi_N =$$

$$\frac{1}{N! \sqrt{2\pi} \prod_{j=1}^N \sigma_{\Phi_j}} \exp\left(-\frac{1}{2} \sum_{j=1}^N \frac{(\Phi_j)^2}{\sigma_{\Phi_j}^2}\right) \prod_{j=1}^N d\Phi_j$$

En se référant au calcul effectué par R.G. STANFIELD [1], la solution la plus probable est donnée par les équations:

$$x = \frac{1}{(\lambda \mu - \nu^2)^{1/2}} \sum_{j=1}^N P_j \frac{\nu \cos \theta_j - \mu \sin \theta_j}{\sigma_{P_j}}$$

$$y = \frac{1}{(\lambda \mu - \nu^2)^{1/2}} \sum_{j=1}^N P_j \frac{\lambda \cos \theta_j - \nu \sin \theta_j}{\sigma_{P_j}}$$

avec

$$\lambda = \sum_{j=1}^N \frac{\sin^2 \theta_j}{(\sigma_{P_j})^2}$$

$$\mu = \sum_{j=1}^N \frac{\cos^2 \theta_j}{(\sigma_{P_j})^2}$$

$$\nu = \sum_{j=1}^N \frac{\sin \theta_j \cos \theta_j}{(\sigma_{P_j})^2}$$

## ANNEXE 2 : MODELES DE COMMUNICATION

Le modèle de communication employé représente des communications parlées. Les communications sont point à point. A l'initialisation, le positionnement temporel des communications est supposé aléatoire dans un domaine T.

Le déroulement d'une communication répond à une distribution des alternats suivant le schéma suivant:

- . 4 alternats de procédure de début de communication
- . N alternats utiles
- . 2 alternats de procédure de fin de communication.

Les alternats de procédure sont de durée (1 + x) secondes, x tiré suivant une loi de rayleigh d'écart-type donné. Les blancs d'alternats sont tirés suivant une loi de rayleigh d'écart-type 1 seconde.

Les alternats utiles sont tirés suivant une loi de rayleigh d'écart-type 4 secondes (60% des cas) ou 8 secondes (40% des cas). Statistiquement la durée d'une communication variera de 50 secondes (alternats courts) à 70 secondes (alternats longs).

Entre 2 communications, on définit un état de veille défini par un temps T résultat de la somme d'une durée minimale (T<sub>0</sub> = 20 secondes) et d'un tirage suivant une loi de rayleigh d'écart-type 24 secondes. Ceci correspond à un taux d'activité des réseaux de 50 %.

## DISCUSSION

### R. ROSE

This system is truly impressive being able to prosecute so many targets in such a small battlefield area. How do you define location? Do you mean just the existence of energy or do you produce an actual location, which is the real end product in battlefield radiolocation?

### AUTHOR'S REPLY

Une mesure goniométrique est effectuée à chaque détection d'énergie. La localisation étant effectuée sur les émissions reconnues, nous produisons une image de l'activité sur le champs de bataille; cette image est renouvelée en temps réel pour suivre les activités des émetteurs. A chaque localisation, nous avons une estimation de la position des émetteurs actifs.

*The direction-finding measurement is made each time energy is detected. As location is performed on recognized emissions, we produce an image of battlefield activity; this image is refreshed in real time so as to monitor emitter activity. Each time we make a location we have an estimate of the position of active emitters.*

### D. YAVUZ

If you have more than one net frequency hopper how do you (or do you attempt to) differentiate between different hop-sets?

### AUTHOR'S REPLY

Le système ne permet pas de mesurer la durée de palier FH. Aussi nous ne pourrions pas distinguer les réseaux.

*The system does not allow measurement of the duration of the frequency-hopper step. Neither can we distinguish the networks.*

### E. KOELBLE

When your system is fast scanning over the frequency band, how great is the false alarm probability for energy detection as a function of the sweep rate and for the signal to noise ratio?

### AUTHOR'S REPLY

La simulation ne prend pas en compte les problèmes liés aux fausses alarmes. Un seuil du rapport signal/bruit à 6 db a été imposé à la détection pour limiter le nombre de fausses détections.

*The simulation does not take false alarm related problems into account. A signal to noise ratio of 6 dB has been imposed on detection in order to limit the amount of false detections.*

### E. LISEC

What hop-rate of true VHF emitters did you assume? Is it related the PR4-G radio?

### AUTHOR'S REPLY

Les émetteurs considérés sont les émetteurs typiques de la bande VHF, c'est à dire entre 100 et 300 sauts par seconde. Les simulations prennent en compte un modèle du TRC 950 de THOMSON-CSF; tous les modes de distribution des émissions EVF n'ont pas été simulés (systèmes à multi-sous-bandes présents dans le PR4G)

*The emitters considered are typical VHF band emitters, i.e., between 100 and 300 hops per second. The simulations assume a THOMSON-CSF TRC 950 emitter; all the EVF emission distribution modes were not simulated (multi-sub-band systems in the PR4-G).*

# RADIO LOCATION THROUGH HIGH RESOLUTION EIGENSTRUCTURE PROCESSING TECHNIQUES THAT YIELD ACCURATE MULTIPATH AOA AND DIFFERENTIAL TIME DELAY ESTIMATES

Leon M. Lewandowski, and  
Keith A. Struckman  
Lockheed Sanders, Inc.  
P.O. Box 868, Nashua, NH 03061

## ABSTRACT

An eigenstructure based, doubly adaptive array processing algorithm is described. This algorithm resolves the differential time of arrival (DTOA) between the direct signal and a delayed replica that arrives over a different ray path. Simulations, using theoretical responses of a circular eight element dipole array, illustrate the resolving characteristics of the algorithm.

## I. INTRODUCTION

HF communication signals often arrive at the receiving antenna over multiple paths. A classic multipath situation of this type is a direct surface wave received in combination with an ionospherically propagated skywave. Resolution of these two components along with an estimate of their differential time delay can be exploited to determine location of the transmitter. This paper describes an eigenstructure based, doubly adaptive processing algorithm which yields an optimum estimate of the differential time of arrival. Angle of arrival (AOA) values for the two ray paths can be calculated by applying superresolution processing algorithms to the array correlation matrix. Digitally beamformed data streams are generated by linearly weighting and additively combining two eigenbeam data streams. The weights are adaptively adjusted to sort the data streams into an interference free copy of the direct path signal and a similarly interference free copy of the delayed path signal. A doubly adaptive algorithm is applied to achieve this sorted signal conditions. Adaptive weights are calculated to maximize a normalized correlation function  $[R_{corr}(t)]^2$  (the ratio of two quadratic forms) as a function of the advance added to the delayed signal. The advance that maximized the normalized correlation function yields an optimum estimate of the differential delay between the two signal propagation paths.

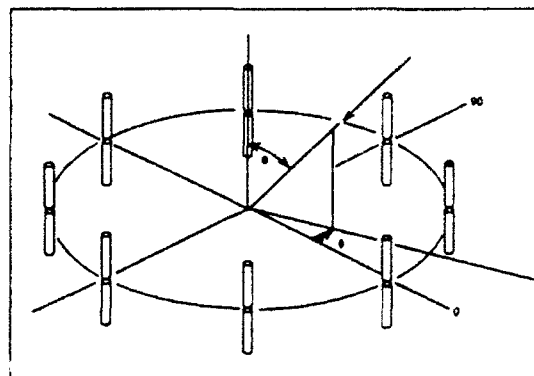
## II. THEORETICAL MULTIPATH PROBLEM

To demonstrate doubly adapted signal resolution, a problem is established by theoretically propagating an incident direct path signal and its delayed path replica onto an idealized eight element dipole array as shown in Figure 1. The voltage recorded at the  $i$ th feedpoint of this eight element circular array, as it intercepts these vertically polarized narrowband planewave signals, is expressed analytically as:

$$V_i(t) = A_i(\theta_{dir}, \phi_{dir})S_{dir}(t) + A_i(\theta_{del}, \phi_{del})S_{del}(t) + N_i(t) \quad (1)$$

05-92-EQ1

where,  $S_{dir}(t)$  is the signal associated with the direct planewave,  $S_{del}(t)$  is the signal associated with the delayed planewave.  $N_i(t)$  is the additive noise at the  $i$ th dipole and  $A_i(\theta, \phi)$  is the response of this dipole in the direction  $(\theta, \phi)$ . Electrically short omni dipole patterns are assumed with responses



CIRCULAR DIPOLE ARRAY Fig 1.

described by:

$$A_i(\theta, \phi) = \sin(\theta) \cdot \{\exp[j2\pi p_a(\omega/c) \cos(\theta - \phi_i)]\} \quad (2)$$

05-92-EQ2

where  $(p\alpha_1 \pi/2, \phi_1)$  are the spherical coordinates of the center of the  $i$ th dipole. The total array response can be written in compact vector notation as:

$$[V(t)] = [A][S(t)] + [N(t)] \quad (3)$$

where:

$$[S(t)] = [S_{dir}(t), S_{del}(t)]^T$$

$$[A] = [A_c(\theta_{dir}, \phi_{dir}), A_c(\theta_{del}, \phi_{del})]$$

05-92-EQ3

and  $A_c(\theta, \phi)$  is an eight element column (steering) vector, each element as described by (2). The array correlation matrix  $[R]$ , experimentally developed by time domain integration over all  $V_i(t)V_j(t)$  products, is given by

$$[R] = [A][S][A]^H + \sigma^2[I] \quad (4)$$

05-92-EQ4

where  $\dagger$  denotes the conjugate transpose and  $[S] = E\{[S(t)][S(t)]^\dagger\}$ .  $\sigma^2[I]$  represents integrated noise voltage contribution to the array correlation matrix. The noise voltages are assumed to be uncorrelated with the signals and uncorrelated among themselves and to have identical variances. A randomly modulated signal, that arrives at the array with different time delays and separate directions, will generate an array correlation matrix that exhibits two signal eigenvectors, i.e., eigenvectors associated with eigenvalues having  $\lambda_i > \sigma^2$ . To identify the structure of these signal eigenvectors, the eigenvector equation:

$$[R][e_k] = \lambda_k [e_k] \quad (5)$$

05-92-EQ6

is recast as:

$$[e_k] = \frac{1}{\lambda_k - \sigma^2} [A] \{ [SA^T][e_k] \} \quad (6)$$

05-92-EQ6

Each signal eigenvector is therefore given by the linear combination of two steering vectors in the form:

$$[e_k] = C_k(\text{dir})[A_c(\theta_{\text{dir}}, \sigma_{\text{dir}})] + C_k(\text{del})[A_c(\theta_{\text{del}}, \phi_{\text{del}})] \quad (7)$$

where:

$$C_k(\text{dir}) = \frac{1}{\lambda_k - \sigma^2} \{ E[S_{\text{dir}}(t)S_{\text{dir}}^*(t)] [A_c(\theta_{\text{dir}}, \phi_{\text{dir}})]^T [e_k] + E[S_{\text{del}}(t)S_{\text{del}}^*(t)] [A_c(\theta_{\text{del}}, \phi_{\text{del}})]^T [e_k] \}$$

$$C_k(\text{del}) = \frac{1}{\lambda_k - \sigma^2} \{ E[S_{\text{dir}}^*(t)S_{\text{del}}(t)] [A_c(\theta_{\text{dir}}, \phi_{\text{del}})]^T [e_k] + E[S_{\text{del}}(t)S_{\text{del}}^*(t)] [A_c(\theta_{\text{del}}, \phi_{\text{del}})]^T [e_k] \}$$

05-92-EQ7

### III. DTOA EQUATION AND WEIGHT OPTIMIZATION TECHNIQUE

The two signal eigenvectors described by (7) span the total steering vector space and are the key components in the resolution solution. The resolution process proceeds by using these two vectors to generate two eigenvector based data streams:

$$S_{\text{sig}}(1, t) = [e_1]^T [V(t)] \quad (8)$$

05-92-EQ8

$$S_{\text{sig}}(2, t) = [e_2]^T [V(t)] \quad (9)$$

05-92-EQ9

In the absence of noise, these data streams span a signal energy space defined as:

$$E[S_{\text{sig}}(1, t)S_{\text{sig}}^*(1, t) + S_{\text{sig}}(2, t)S_{\text{sig}}^*(2, t)] = \lambda_1 + \lambda_2 \quad (10)$$

05-92-EQ10

A simple analysis, based on (7), (8) and (9), shows that these two data streams can be weighted and summed to eliminate either direct path or delayed path signals. Assume that these weights have been calculated by some scheme and result in a direct signal data stream:

$$U_{\text{dir}}(t) = W_{\text{dir}}(1)S_{\text{sig}}(1, t) + W_{\text{dir}}(2)S_{\text{sig}}(2, t) \quad (11)$$

05-92-EQ11

that contains the direct signal and weighted noise voltages but zero delayed signal components. Similarly a delayed signal data stream:

$$U_{\text{del}}(t) = W_{\text{del}}(1)S_{\text{sig}}(1, t) + W_{\text{del}}(2)S_{\text{sig}}(2, t) \quad (12)$$

05-92-EQ12

is assumed that does not contain direct signal components.

Differential time delay between these two signals can be computed by observing the correlation function:

$$|R_{\text{corr}}(\tau)|^2 = \frac{|E[U_{\text{dir}}(t)U_{\text{del}}^*(t + \tau)]|^2}{E[U_{\text{dir}}(t)U_{\text{dir}}^*(t)]E[U_{\text{del}}(t + \tau)U_{\text{del}}^*(t + \tau)]} \quad (13)$$

05-92-EQ13

as the delayed data stream is numerically advanced. The value of tau that maximizes this function is exactly equal to the true multipath delay if the signal-to-noise ratio (SNR) is equal to  $\infty$ .

DTOA equation (13) is a maximization function that depends on precomputed weights to sort the signals prior to time delay scanning. An alternative to this difficult precomputation task is to include weight computation within the correlation maximization process. Sets of new correlation maximization weights are computed at each  $\tau_{\text{test}}$  step in the correlation scanning process. At any general  $\tau_{\text{test}}$  point, these optimized weights will maximize the correlation but in general do not sort the signals. However, when  $\tau_{\text{test}} = \tau_{\text{true}}$ , the signals will be sorted since only this condition will ensure, at  $SNR = \infty$ , a maximized correlation function. The resolution process described herein is based on the ability to calculate, in closed form, weights that maximize the correlation function. The resolution process described herein is based on the ability to calculate, in closed form, weights that maximize the correlation function. Weight optimization is described as a two step process, direct signal weights are optimized followed by delayed signal weight optimization.

First, it is assumed that accurate delayed signal weights have been precomputed. The correlation function is then given by an equation that is the ratio of two quadratic forms:

$$|R_{\text{corr}}(\tau)|^2 = \frac{[W_{\text{dir}}(1) W_{\text{dir}}(2)] [M_1(1,1) M_1(1,2)] [W_{\text{del}}(1)]}{[M_1(1,1) M_1(2,2)] [W_{\text{del}}(2)]}$$

$$DT_1 = \frac{[W_{\text{dir}}(1) W_{\text{del}}(2)] [D_1(1,1) D_1(1,2)] [W_{\text{del}}(1)]}{[D_1(1,1) D_1(2,2)] [W_{\text{del}}(2)]} \quad (14)$$

where:

$$D_1(1,1) = E[S_{\text{sig}}(1, \eta)S_{\text{sig}}^*(1, \eta)]$$

$$D_1(1,2) = E[S_{\text{sig}}(1, \eta)S_{\text{sig}}^*(2, \eta)]$$

$$D_1(2,2) = E[S_{\text{sig}}(2, \eta)S_{\text{sig}}^*(2, \eta)]$$

$$M_1(1,1) = |E[S_{\text{sig}}(1, \eta)U_{\text{del}}^*(1, t + \tau)]|^2$$

$$M_1(1,2) = E[S_{\text{sig}}(1, \eta)U_{\text{del}}^*(1, t + \tau)]E[S_{\text{sig}}^*(2, \eta)U_{\text{del}}(1, t + \tau)]$$

$$M_1(2,2) = |E[S_{\text{sig}}(2, \eta)U_{\text{del}}^*(1, t + \tau)]|^2$$

$$DT_1 = E[U_{\text{del}}(1, t + \tau)U_{\text{del}}^*(1, t + \tau)]$$

05-92-EQ14

The object is to maximize equation (14) over the direct signal weights. This apparently difficult problem, is instead extremely simple, since the maximization ratio of quadratic forms is easily computed [2]. The maximum value over the weights is given by largest root of the determinant:

$$\begin{vmatrix} M_1(1,1) & M_1(1,2) \\ M_1(1,2) & M_1(2,2) \end{vmatrix} - \lambda \begin{vmatrix} D_1(1,1) & D_1(1,2) \\ D_1(1,2) & D_1(2,2) \end{vmatrix} = 0 \quad (15)$$

05-92-EQ15

divided by the function  $DT_1(t)$ . DTOA is calculated by scanning the ratio of the two functions  $\lambda(t)/DT_1(t)$  and searching for a maximum. This maximum value can then be inserted into equation (14) and solved to yield the direct signal weights. Specific direct signal weight values need not be calculated if DTOA resolution is the single requirement. At this stage in the DTOA computational process, a significant problem still exists, that of precomputing the delayed signal weights.

Step two solves the delayed signal weight computation problem by an algebraic expansion of the only nonzero root of equation (15), which is:

$$\lambda_{\max} = \frac{1}{DT_2} \{D_1(1,1)M_1(2,2) + D_1(2,2)M_1(1,1) - [D_1(1,2)M_1(1,2) + D_1(1,2)M_1(1,2)']\} \quad (16)$$

$$\text{where, } DT_2 = D_1(1,1)D_1(2,2) - |D_1(1,2)|^2$$

05-92-EQ16

This root is a function of the precomputed delayed signal weights. When these weights are included as algebraic terms instead of numerical values and the result divided by the  $DT_1$  (also a function of delayed signal weights, an interesting ratio is observed. This new function is also a ratio of quadratic forms given by:

$$|R_{\text{con}}(t)|^2 = \quad (17)$$

$$DT_1 = \frac{[W_{\text{del}}(1) W_{\text{del}}(2)] \begin{bmatrix} M_2(1,1) & M_2(1,2) \\ M_2(1,2) & M_2(2,2) \end{bmatrix} \begin{bmatrix} W_{\text{del}}(1) \\ W_{\text{del}}(2) \end{bmatrix}}{\begin{bmatrix} W_{\text{del}}(1) & W_{\text{del}}(2) \end{bmatrix} \begin{bmatrix} D_2(1,1) & D_2(1,2) \\ D_2(1,2) & D_2(2,2) \end{bmatrix} \begin{bmatrix} W_{\text{del}}(1) \\ W_{\text{del}}(2) \end{bmatrix}}$$

05-92-EQ17

None of the  $M_2...$  or  $D_2(,)$ ... terms presented in equation (17) involve weights, therefore this second adaptive equation is maximized over delayed signal weights by calculating the maximum eigenvalue of a second determinant which has a form similar to equation (15). The terms required for this calculation will not be included since they are readily derived.  $D_2(i,j)$  terms have the same functional dependency as  $D_1(i,f)$  with  $t$  replaced by  $t + \tau$ . For reference purposes,  $M_2(1,2)$  is given by:

$$M_2(1,2) = Z(1,1)Z(1,2)D(2,2) + Z(2,2)D(1,1) - [Z(1,2)Z(2,1)D(1,2) - Z(1,1)Z(2,2)D(1,2)]$$

where

$$\begin{aligned} Z(1,1) &= E[S_{\text{sig}}^*(1, t) S_{\text{sig}}(1, t + \tau)] \\ Z(1,2) &= E[S_{\text{sig}}^*(1, t) S_{\text{sig}}(2, t + \tau)] \\ Z(2,1) &= E[S_{\text{sig}}^*(2, t) S_{\text{sig}}(1, t + \tau)] \\ Z(2,2) &= E[S_{\text{sig}}^*(2, t) S_{\text{sig}}(2, t + \tau)] \end{aligned} \quad (18)$$

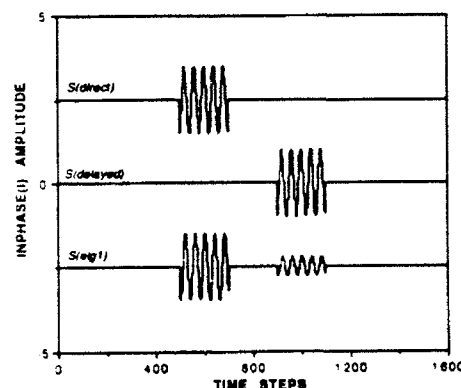
05-92-EQ18

Determinant number two generally exhibits two nonzero eigenvalues which are functions of time delay. DTOA solutions are obtained by scanning the ratio, maximum eigenvalue divided by  $DT_2$ . This simple calculation is all that is necessary for determining DTOA values. If signal sorting is desired, the two sets of weights can be calculated by reverse

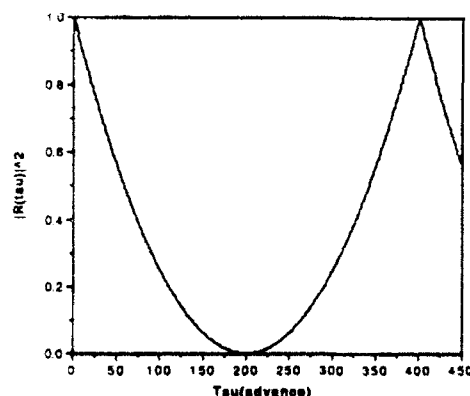
processing i.e., correlation maximized over time delay, which is then used for delayed signal weight calculations followed by direct signal weight calculations. It should be noted that the maximization function is a ratio, therefore either set of weights can be multiplied by an arbitrary nonzero complex scalar without changing the value. This indeterminant problem does not, however, prevent unique signal resolution.

#### IV. SIMULATION RESULTS

A set of simulations are presented to illustrate the performance of the doubly adaptive resolution technique. A simple pulse modulated CW analytic signal is uniformly sampled at 40 pts. per cycle to generate an incident signal. The in phase (1) component of this pulse is shown on the upper trace, Figure 2.



S(t) PATTERNS (delay = 400) Fig. 2



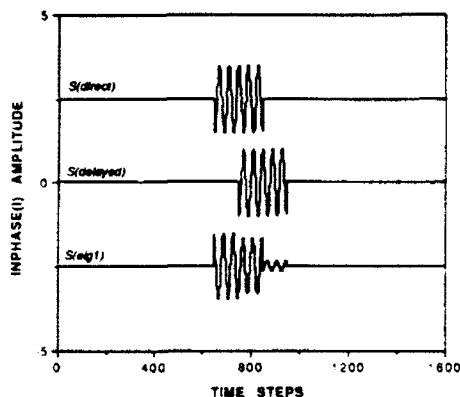
CORRELATION CURVE (delay = 400) Fig. 3

The first example is based on an equal amplitude, significantly delayed signal (three times the pulse width), as shown on the middle trace. Incident angles in the first quadrant,  $(\Theta_{\text{dir}}, \Phi_{\text{dir}}) = 90., 45.)$ ,  $(\Theta_{\text{del}}, \Phi_{\text{del}}) = 30., 45.)$  are assumed. A covariance matrix having a ratio of eigenvalues  $\lambda_2/\lambda_1 = .170$  was generated by approximating the expectation function by time domain integration over the window shown. Eigenvector data stream  $S_{\text{sig}}(1, t)$  associated with the largest eigenvalue is normalized and displayed on the lower trace of the figure. Computations based on the doubly adaptive correlation function (17) exhibits (Figure 3), two correlation values equal to 1.0, the first a false peak at time step 0 and a second triangular peak, at the correct time step = 400. This simple pulse problem could have been solved by a simple clip and scan technique, but its use in correlation function processing clearly demonstrates

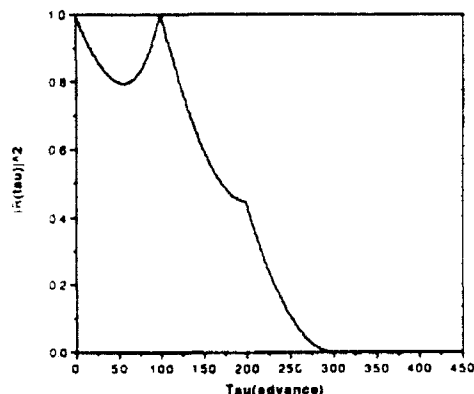


the doubly adaptive concept. It should be noted that a unity magnitude peak at zero time is always generated by the correlation maximization process described herein.

The second example, is based on parameters equal to the first example except for amount of signal delay, which was set to equal the pulse width divided by two as shown in Figure 4. Here, the eigenvector data stream  $S_{eig}(1,t)$  exhibits the interaction between the direct and delayed signals, however, the correlation peak  $|R_{corr}(\tau)|^2 = 1.08\tau_{test} = 100$ , is well defined and at the correct delay as shown in Figure 5. The eigenvalue ratio for this delay equals 0.206. this ratio is reduced to 0.061 at a delay equal to 1/10 the pulse width. At zero delay, the second eigenvalue is equal to the noise variance.



S(i) PATTERNS (delay = 100) Fig. 4



CORRELATION CURVE (delay = 100) Fig. 5

The time delay resolution of the signals described in the above examples is based on an internal process that spatially notches each signal. Direct signal resolution, requires a null at (30., 45.), resolution of the delayed signal required a null at (40., 45.). Example 2 parameters are used to demonstrate the adaptive beamformed patterns. Figure 6(a) shows the beamformer output power as a function of signal arrival angle, within the first quadrant, using weights designed to reject the delayed signal. Figure 6(b) shows the power pattern associated with the beamformer designed to receive the delayed signal and reject the direct signal. Obviously, the beamformed patterns exhibit the correct response, peaks in desired direction and nulls in the desired cancellation direction. The process drivers are the nulls, not the peaks.

These beamformers were based on weighted and summed signal eigenvectors, using weights calculated to satisfy equations 17 and 14.

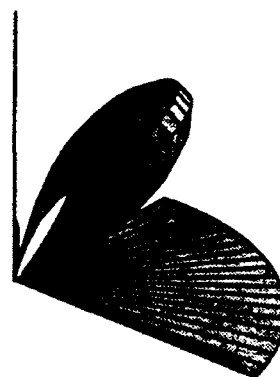
## V. DISCUSSION

The doubly adaptive concept provides a straightforward solution to the twopath DTOA resolution problem. Expanded eigenstructure techniques that yield solutions to a larger class of problems are being investigated. Experiments using the doubly adaptive concept described herein, but minus the eigenstructure processing was used to sort direct path signals and ionospherically refracted signals. Differential time delays of approximately 1.9 milliseconds were observed for both pulse and phase modulated rf signals.

### BEAMFORMED POWER PATTERNS



Copy Beam(direct signal) Fig. 6(a)



Copy Beam(delayed signal) Fig. 6(b)

## VI. REFERENCES

- [1] T.J. Shan, M. Wax and T. Kailath, "On spatial smoothing for direction-of-arrival estimation of coherent signals," IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-33, pp. 806-811, Aug. 1985.
- [2] F. R. Gantmacher, The Theory Of Matrices, Vol. 1, New York: Chelsea Publishing Co., 1977, Ch. 10.

## DISCUSSION

### C. GOUTELARD

Quelles sont vos possibilités de discrimination sur les trajets multiples: nombre, séparation temporelle?

*What are your discrimination capabilities on multiple paths: number, temporal separation?*

### AUTHOR'S REPLY

The number of combinations grow combinatorially, e.g., 3 modes lead to 3 interpath delays, 4 modes to 6 interpath delays, etc. The issue is even if the higher order multipath modes could be resolved in terms of multipath delays, would it be possible to infer origin of the source depending, of course, on how well ionospheric propagation phenomenology (in real time) can be known and used to invert the geometry to locate the source.

### J. BELROSE

I suggest that you should include propagation parameters to make your computer simulation more realistic. You spoke about distances of 300 and 800 km. Dependent on frequency and distance, the ground wave could be relatively weak, compared with the skywave, even for propagation over sea water.

### AUTHOR'S REPLY

This is true: the lower frequencies (say  $< 10$  MHz) will tend to propagate relatively long distances over sea water ( $\sigma = 5$  mhos/meter, dielectric constant 80). But at these frequencies (well below the oblique incidence MUF) we'll see multiple skywave paths. This underlines Professor Goutelard's question regarding the potential of our processing techniques' ability to resolve multiple paths (i.e., more than two). The method is generalizable to more than two paths, but just as my presentation showed there are limits to resolvability that depend on array size and especially on the spectral aperture of the signal, and here I mean specifically the root mean square bandwidth of the signal spectrum. As is well known from parameter estimation theory, the resolvability (actually more correctly the accuracy of the estimate) is inversely proportional to the RMS bandwidth. The cases would need to be evaluated for their potential depending on S/N, signal bandwidth, and propagation geometry, on a case by case basis.

# Multiple Emitter Direction Finding from a Circular Array by Single Snapshot Processing

Y. Tanık

A. (Tuncay) Koç

Middle East Technical University  
Electrical and Electronic Engineering Department  
İnönü Bulvarı, 06531 Ankara  
Turkey

## 1. SUMMARY

In this work, a new algorithm for multiple emitter direction finding for circular arrays is proposed. The algorithm uses a single snapshot for processing. The observation is first transformed into frequency domain. Then, by using the Fourier series coefficients of pattern of the sensors, it is transformed into a suitable sequence. The linear prediction method is applied to the resulting sequence to determine the direction of arrivals from the zeros of corresponding linear prediction filter. An iterative correction scheme is used to improve the accuracy of the estimates. Computer simulations have been performed to evaluate the performance of the algorithm. It has been observed that the algorithm operates successfully if the array circumference is comparable to the wavelength.

## 2. INTRODUCTION

High-resolution direction-of-arrival (DOA) estimation is important in many sensor systems such as radar, sonar, electronic surveillance, and seismic exploration. In the simplest case, the signals received by the sensors consist of scaled and delayed replicas of the waveform radiated by a single source. In a more realistic and complicated scenario, there may be multiple sources and multiple propagation paths from the sources to the sensors.

The DOA estimation problem has received considerable attention in the last 40 years, and a variety of techniques for its solution have been proposed. The number of snapshots (observation vectors at the outputs of array sensors) required for processing can be used for classifying these techniques. The algorithms that can use only one observation vector may be termed as single snapshot methods while those that require many observations may be called multiple snapshot methods.

Single snapshot methods can be seen to be of advantage where the observation time should be small as compared to the incoherency time of the incoming signals. Furthermore, the complications that arise in data acquisition, storage and computation due to increased number of snapshots, favor the selection of single snapshot methods, especially in real-time

applications.

On the other hand, multiple snapshot methods are based on the averaging of the observations. During the relatively long observation time, the only parameters that remain unchanged are the DOAs of the signals, while the other parameters can change significantly. Therefore, a kind of separation between DOAs and other parameters of the signals becomes possible due to the time averaging.

Among the algorithms that can use a single snapshot, the maximum likelihood (ML) technique, Burg's maximum entropy (ME) method, and linear prediction (LP) modeling techniques can be mentioned.

The ML technique is known to be an optimal technique and its application to the DOA estimation problem has been extensively studied. It has the advantage that it can be applied to the case of coherent signals and there is no restriction on the array geometry. The application of ML estimation requires a highly nonlinear optimization procedure over a multidimensional parameter space. Several optimization techniques have been proposed all suffering from high computational cost and there is no guarantee for global convergence. Recently, two algorithms have been proposed to overcome the associated numerical difficulties, namely, the method of alternating projections [1], and the expectation-maximization (EM) algorithm [2], however, global convergence is not guaranteed in these algorithms either.

The ME method, [3], is often successful and widely used, but has certain limitations, especially bias and sensitivity in parameter estimates. The parameter estimates of the ME method are significantly biased even in the presence of nearly perfect observation, [4].

The LP modeling techniques, [5], were also adopted for the array problem and stirred much interest in the so-called high-resolution estimation procedures, [6]. Model-based approaches have the distinct advantage of producing high-resolution estimates under extremely small data sample sizes. These

approaches are largely restricted to uniform linear array geometry, [7], but it is known that in most cases full peripheral coverage in direction finding (DF) is required, and circular arrays rather than linear arrays are desirable because of their more compact structures and the inherent symmetry for 360° coverage.

The eigenstructure based methods are multiple snapshot methods since they require a large number of snapshots for DOA estimation, and their performances degrade for small number of snapshots. Most of the eigenstructure based methods first form an estimate of the covariance matrix by using many observation vectors, then decompose the observed covariance matrix into two orthogonal spaces, commonly referred to as the signal and noise subspaces, and estimate the DOAs from one of these spaces. The more frequently referenced techniques are MUSIC [8], MIN-NORM [9], [10], and ESPRIT [11]. The conventional MUSIC does not have any restriction on array geometry, and its performance advantages are substantial, but they are achieved at a considerable cost in computation (searching over parameter space) and storage (of array calibration data). ESPRIT was introduced to reduce these computation and storage costs; but the reductions are achieved by requiring that the sensor array possess a displacement invariance, i.e., sensors occur in matched pairs with identical displacement vectors. Although generally applicable, MIN-NORM was developed for arrays of linear equispaced matched omnidirectional sensors.

The eigenstructure based methods are known to yield high resolution and asymptotically unbiased estimates, even in the case that the sources are partially correlated. Theoretically, these methods encounter difficulties only when the signals are perfectly correlated. In practice however, these algorithms fail even when the signals are highly correlated, as happens, for example, in multipath propagation or in military scenarios involving smart jammers.

As a solution to this problem, the idea of spatial smoothing, has been suggested by Evans et al. [12] and further developed by Shan et al. [13], and was extensively studied by Williams et al. [14], and Pillai and Kwon [15]. The spatial smoothing technique is based on a preprocessing scheme that partitions the total array of sensors into subarrays and then generates the average of the subarray output covariance matrices, and therefore is restricted to arrays containing translational equivalent subarrays such as uniform linear arrays, [16]. The averaging process essentially decorrelates the signals and thus coherent signals can be resolved by the application of any of the eigenstructure based algorithms. A disadvantage of the spatial smoothing algorithm is that it significantly reduces the effective array aperture which

results in reduction of the resolving power.

In this work, we propose an algorithm for multiple emitter DF employing a uniform circular array which uses a single snapshot of data from the array. This is an advantage since most of the single snapshot algorithms that have been proposed are either restricted in the array structure or they have convergence problems. For example, LP methods have only been applied to uniform linear arrays and ML method requires an optimization over a highly non-linear surface.

The proposed algorithm provides estimates for the DOAs and complex envelopes of the incoming plane waves in an iterative fashion. First a transformation is applied on the observed data and then a model-based approach is used to obtain initial DOA estimates, which are used to estimate complex envelopes which are in turn used for improving the DOA estimates. Since the algorithm is based on single snapshot processing, it has no restriction on the coherency of the sources.

The paper is organized as follows. In Section 3, we formulate the problem and present the proposed algorithm. Computer simulations that demonstrate the performance of the algorithm are presented in Section 4. Our concluding remarks are given in Section 5.

### 3. PROBLEM FORMULATION

The array under consideration is a uniform circular array composed of  $M$  identical sensors. Assume that  $L$  ( $L < M$ ) narrow-band plane waves, at a known wavelength  $\lambda$ , impinge on the array from directions  $\theta_1, \theta_2, \dots, \theta_L$  with respect to the line of reference of the array. Let  $r$  represent the array radius and let  $\phi_k$  be the angular position of the  $k^{\text{th}}$  sensor measured from the array reference. The  $k^{\text{th}}$  sensor is located at

$$\phi_k = \frac{2\pi}{M}k, \quad \text{for } k: 0, \dots, M-1. \quad (1)$$

The snapshot at the  $k^{\text{th}}$  sensor is

$$y_k = s_k + u_k, \quad \text{for } k: 0, \dots, M-1, \quad (2)$$

where  $s_k$  and  $u_k$  are the signal and noise samples at the  $k^{\text{th}}$  sensor, respectively.

The following assumptions are made on the described model and are considered valid throughout the paper:

- The sensor array is located at the far field such that planar wave expressions can be used for incident waves.
- The array sensor noise is zero-mean Gaussian. In addition, the noise samples are uncorrelated

between themselves, and have identical variances,  $\sigma^2$ .

- All incident waves are assumed to impinge upon the array at  $0^\circ$  elevation. Therefore, all signal processing algorithms are constrained to estimate the DOA parameters in azimuth only.

Taking the center of the circle as the reference point of phase, and using superposition, the signal received by the  $k^{\text{th}}$  sensor can be expressed as

$$s_k = \sum_{l=1}^L a_l P(\theta_l - \frac{2\pi}{M}k) \exp \left\{ j \frac{2\pi r}{\lambda} \cos(\theta_l - \frac{2\pi}{M}k) \right\}, \quad (3)$$

for  $k : 0, \dots, M-1$ , where  $a_l$  is the complex envelope of the  $l^{\text{th}}$  plane wave, and  $P(\cdot)$  denotes the common pattern of the antennas. When the sensors are omnidirectional,  $P(\cdot)$  will be a constant.

Defining

$$R(\theta) \triangleq P(\theta) \exp \{ j\beta \cos(\theta) \}, \quad (4)$$

where

$$\beta \triangleq \frac{2\pi r}{\lambda}, \quad (5)$$

(3) can be rewritten as

$$s_k = \sum_{l=1}^L a_l R(\theta_l - \frac{2\pi}{M}k), \quad \text{for } k : 0, \dots, M-1. \quad (6)$$

Since  $R(\theta)$  is periodic with period  $2\pi$ , it can be expanded into a Fourier series (FS) as

$$R(\theta) = \sum_{n=-\infty}^{\infty} r_n e^{jn\theta}. \quad (7)$$

where  $r_n$  are the FS coefficients.

Substituting (7) in (6),

$$\begin{aligned} s_k &= \sum_{l=1}^L a_l \sum_{n=-\infty}^{\infty} r_n e^{jn(\theta_l - \frac{2\pi}{M}k)} \\ &= \sum_{n=-\infty}^{\infty} r_n e^{-j\frac{2\pi}{M}nk} \sum_{l=1}^L a_l e^{jn\theta_l}, \end{aligned} \quad (8)$$

for  $k : 0, \dots, M-1$ .

By expressing the received signals in this form, the terms possessing the information concerning the DOAs,  $\theta_l$ , and the complex envelopes,  $a_l$ , of the incident plane waves are separated from the terms that contain the information on the array structure. To simplify the notation, let us define

$$\alpha_n \triangleq \sum_{l=1}^L a_l e^{jn\theta_l}, \quad (9)$$

and

$$\beta_n(k) \triangleq r_n e^{-j\frac{2\pi}{M}nk}. \quad (10)$$

Thus, (8) can be reexpressed as

$$s_k = \sum_{n=-\infty}^{\infty} \alpha_n \beta_n(k), \quad \text{for } k : 0, \dots, M-1. \quad (11)$$

Since the array structure is known,  $\beta_n(k)$  are known values and the unknown parameters of concern are included in  $\alpha_n$ . Obviously, only the noisy data,  $\{y_k\}$ , given in (2) is available. For the moment we will omit the noise; it will be taken into account in the sequel.

Equation (9) is in a well-known form and estimation of  $\{\theta_l\}$  and  $\{a_l\}$  from  $\{\alpha_n\}$  has been extensively studied in the literature. Several methods have been used and a detailed survey can be found in [17] and [18]. One of the most popular techniques that have been used is the LP method. Ideally, the LP filter which filters  $\{\alpha_n\}$  will have zeros at DOAs. However,  $\alpha_n$  are unknown and estimation of  $\alpha_n$  from the observed data is necessary.

The sequence  $\{\alpha_n\}$  should be determined from (11). Apparently, this equation contains an infinite number of unknowns whereas only  $M$  equations are available. Therefore,  $\{\alpha_n\}$  cannot be determined in a straightforward fashion. In order to overcome this problem to an extent, we decompose the infinite summation in (11) into blocks of length  $M$  as

$$\begin{aligned} s_k &= \sum_{q=-\infty}^{\infty} \sum_{n=-\frac{M}{2}+1}^{\frac{M}{2}} \alpha_{qM+n} \beta_{qM+n}(k), \\ &= \sum_{n=-\frac{M}{2}+1}^{\frac{M}{2}} \beta_n(k) \sum_{q=-\infty}^{\infty} \frac{r_{qM+n}}{r_n} \alpha_{qM+n}, \end{aligned} \quad (12)$$

for  $k : 0, \dots, M-1$ , where we have made use of the following fact:

$$\beta_{qM+n}(k) = \frac{r_{qM+n}}{r_n} \beta_n(k). \quad (13)$$

Defining

$$U_n \triangleq \sum_{q=-\infty}^{\infty} \frac{r_{qM+n}}{r_n} \alpha_{qM+n}, \quad (14)$$

for  $n : (-M/2) + 1, \dots, (M/2)$ , (12) becomes

$$s_k = \sum_{n=-\frac{M}{2}+1}^{\frac{M}{2}} U_n \beta_n(k), \quad \text{for } k : 0, \dots, M-1. \quad (15)$$

From the definition of  $U_n$ , it can be seen that these terms are related to  $\alpha_n$  by

$$U_n = \alpha_n + \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} \frac{r_{qM+n}}{r_n} \alpha_{qM+n}, \quad (16)$$

for  $n : (-M/2) + 1, \dots, (M/2)$ , and defining the second term in the right hand side of (16) as the distortion term,

$$d_n \triangleq \sum_{\substack{q=-\infty \\ q \neq 0}}^{\infty} \frac{r_{qM+n}}{r_n} \alpha_{qM+n}, \quad (17)$$

for  $n : (-M/2) + 1, \dots, (M/2)$ , (16) can be rewritten as

$$U_n = \alpha_n + d_n, \quad \text{for } n : -\frac{M}{2} + 1, \dots, \frac{M}{2}. \quad (18)$$

It would be desirable if the ratios of the FS coefficients appearing in (17),  $r_{qM+n}/r_n$ , were negligibly small for  $q \neq 0$  so that  $d_n \approx 0$ , or equivalently  $U_n \approx \alpha_n$ , since (15) would directly yield  $\alpha_n$  in this case.

In practice, the observations will inevitably be noisy and the estimation of  $\{U_n\}$  is required. We use the minimum mean square error criterion for this estimation. Defining

$$J \triangleq \sum_{k=0}^{M-1} \left| y_k - \sum_{n=-\frac{M}{2}+1}^{\frac{M}{2}} U_n \beta_n(k) \right|^2, \quad (19)$$

and minimizing the mean square error,  $J$ , with respect to  $U_n$ , for  $n : (-M/2) + 1, \dots, (M/2)$ , it can be shown that

$$\hat{U}_n = \frac{Y(-n)}{Mr_n}, \quad \text{for } n : -\frac{M}{2} + 1, \dots, \frac{M}{2}, \quad (20)$$

where  $Y(n)$  is the discrete Fourier transform (DFT) of the observation sequence,  $\{y_k\}$ , and the required values of  $Y(-n)$ , for  $n : (-M/2) + 1, \dots, (M/2)$ , can be found by using the periodicity of the discrete Fourier series.

Note that the solution will be finite and unique if  $|r_n| \neq 0$  for  $n : (-M/2) + 1, \dots, (M/2)$ .

The estimates for  $U_n$ ,  $\hat{U}_n$ , given by (20) will differ from the exact values by additive noise terms as

$$\begin{aligned} \hat{U}_n &= U_n + z_n, \\ &= \alpha_n + d_n + z_n, \end{aligned} \quad (21)$$

for  $n : (-M/2) + 1, \dots, (M/2)$ , where  $z_n$  are the transformed form of the sequence of  $w_k$  and they are zero-mean, uncorrelated with unequal variances,

i.e.,  $E\{z_n\} = 0$  and  $E\{z_n z_m^*\} = [\sigma^2 / (Mr_n r_m^*)] \delta_{n,m}$ . Here  $E\{\cdot\}$  is the expectation operator and  $\delta$  is the Kronecker delta.

Since the estimated values  $U_n$  contain distortions on  $\alpha_n$  in addition to noise terms, it is necessary to find good estimates for  $d_n$  in order to estimate  $\alpha_n$  accurately.

Fortunately, the distortion terms,  $d_n$ , will not be too large since the FS coefficients form a rapidly decreasing sequence for smooth functions. Therefore, the infinite sum over  $q$  in (17) can be truncated at some  $Q$  which can be very small. The distortion terms include  $\alpha_{qM+n}$ , hence  $\theta_l$  and  $a_l$ , and therefore, an iterative estimation algorithm becomes inevitable.

The proposed algorithm can be summarized as follows:

#### The Algorithm:

1. Find  $\hat{U}_n$  using (20).

This corresponds to finding the DFT of the snapshot and then weighting it suitably.

2. Set  $\hat{\alpha}_n = \hat{U}_n$  as initial estimates.

3. Apply a standard, forward or forward-backward LP method to  $\{\hat{\alpha}_n\}$  in order to obtain the predictor coefficients.

Construct the polynomial corresponding to the  $z$ -transform of the LP filter and find the roots of this polynomial.

The angles of these roots are the estimates for  $\theta_l$ ,  $\hat{\theta}_l$ .

4. Find the estimates for  $a_l$ ,  $\hat{a}_l$ , which minimize

$$\sum_{k=0}^{M-1} \left| y_k - \sum_{l=1}^L \hat{a}_l P(\hat{\theta}_l - \phi_k) e^{j\hat{\theta}_l \cos(\hat{\theta}_l - \phi_k)} \right|^2. \quad (22)$$

This procedure can be shown to be equivalent to the ML estimation. It can be seen that  $\hat{a}_l$  are obtained by solving a system of linear equations.

5. Using the estimates  $\hat{\theta}_l$  and  $\hat{a}_l$  in (9) and (17), obtain estimates for  $d_n$ ,  $\hat{d}_n$ . The infinite sum in (17) can be truncated as mentioned previously.

6. Improve  $\hat{\alpha}_n$  by subtracting  $\hat{d}_n$  from  $\hat{U}_n$ , (see equation (18)). Go to step 3 if convergence is not achieved.

The following points related to the algorithm should be mentioned:

- i. The number of emitters is a parameter that should also be determined. Ideally, the order of the LP filter,  $K$ , should be

equal to the number of emitters, however, the number of emitters is not known. Therefore, the highest possible order ( $N/2$  for forward linear predictor (FLP), and  $2N/3$  for forward-backward linear predictor (FBLP), where  $N$  is the data length. [18]) for the LP filter can be chosen. At the end of the algorithm the number of emitters can be estimated by inspecting  $\{a_l\}$ , i.e., if  $|a_l| \approx 0$  for some  $l, l = 1, \dots, K$ , the corresponding signals should be eliminated.

- ii. In general, the FBLP method is preferred to the FLP method, because the FBLP method applies both forward and backward predictions to the same data set and in this manner some extra data points are obtained over which to average the prediction errors. Therefore, the FBLP method yields improved estimates of predictor coefficients as compared to the FLP method. [17], [18]. Furthermore, in our case the data used for prediction has no direction dependency, so the FBLP method is very convenient.
- iii. The LP method is based on the minimization of the sum of the square of linear prediction errors. Actually, a minimization of the weighted sum of these errors is also possible.

One such method was proposed by Nikias and Scott, [19], in which the weights represent the energy of the data used to form the prediction errors.

The linear predictor input noise power is inversely proportional to  $|r_n|^2$ , and if the FS coefficients become small for some  $n$ ,  $n = (-M/2) + 1, \dots, (M/2)$ , then the corresponding noise power will be too high. Therefore,  $|r_n|^2$  can be used as weighting coefficients so that the terms that have more noise will have less effect on the result.

- iv. The truncation of the infinite sum in step 5 of the algorithm should be considered. For symmetric antenna patterns the FS coefficients have the property that

$$r_{-n} = r_n. \quad (23)$$

As a special case, for omnidirectional antennas the FS coefficients are found to be

$$r_n = j^n J_n(\beta) \quad (24)$$

satisfying the property given by (23), where  $J_n(\beta)$  is the Bessel function of the first kind of order  $n$ , and  $\beta$  is as defined by (5).

If  $M > \beta^2$ , due to the following property of Bessel functions

$$J_x(\beta) > J_y(\beta) \text{ if } y > x \text{ when } x \geq \beta,$$

it can be shown that the ratios of the FS coefficients in (17) become negligibly small for sufficiently large values of  $q$ , and the truncation of the infinite sum is justified.

Considering (17), it can be easily seen that for  $q = -1$  the ratio of the FS coefficients which satisfy the property given by (23), becomes unity for  $n = M/2$ , resulting in significant distortion.

Due to this fact, in the application of the LP techniques we prefer to use the sequence of  $\hat{a}_n$ , for  $n = (-M/2) + 1, \dots, (M/2) - 1$ , excluding the information corresponding to  $n = M/2$ .

- v. The estimates of  $\theta_l$  and  $a_l$  are obtained in different steps. This parameter separation simplifies the nonlinear optimization problem by reducing the dimension of the parameter space over which optimization should be performed.
- vi. Simulation results have shown that the algorithm described above performs well in most cases. However, in some instances an oscillatory behavior was observed. In order to prevent this undesirable behavior the following modification was made in step 6:

$$\hat{a}_n = \hat{r}_n - \gamma \hat{d}_n, \quad (25)$$

where  $\gamma$  is chosen to be close to but less than unity.

#### 4. SIMULATION RESULTS

In order to demonstrate the performance of the proposed algorithm under various situations, several computer simulations have been conducted. In all the simulations, eight isotropic sensors,  $M = 8$ , and two narrow-band emitters,  $L = 2$ , were considered.

To observe the effect of the array size on the performance of the algorithm, six different values of  $\beta$ , namely,  $\beta = 0.5, 1, 1.5, 2, 2.5, 3$  were considered. For all these values of  $\beta$ , two different signal-to-noise ratios (SNR), 100 dB and 80 dB were studied. The SNR is defined as  $\text{SNR} = 10 \log(\text{SP}/\sigma^2)$ , where SP is the maximum signal power, i.e.,

$$\text{SP} \triangleq \max_l (|a_l|^2),$$

and  $\sigma^2$  is the average noise power.

In all these cases the two emitters were located at  $0^\circ$  and  $10^\circ$ , and the incoming signals were assumed to arrive with the same phase,  $0^\circ$ , and their magnitudes being 1 and 0.5, respectively. In these simulations 100 runs were performed and the root-mean-square (RMS) error of the estimates for each DOA

were computed. These results are compared to the Cramér-Rao lower bound (CRB), [20], for each case.

In the 3<sup>rd</sup> step of the algorithm, the weighted FBLP method with  $|r_n|^2$  as weighting coefficients, was used. The order of the LP filter was chosen to be equal to the number of emitters, i.e.,  $K = L$ . In the 5<sup>th</sup> step of the algorithm, the infinite sum was truncated at  $Q = 1$ , since the results were not improved by choosing a higher value of  $Q$ . In finding  $\hat{\alpha}_n$ , given by (25),  $\gamma$  was chosen to be 0.9999.

The simulation results are summarized in Table 1 and Table 2 for SNR = 100 dB, and SNR = 80 dB, respectively. The tables include the resulting RMS errors (in degrees) of the estimates of both DOAs along with the number of iterations required for convergence. The CRBs (in degrees) are also given for comparison. The convergence criterion is mainly based on the cost function, given by (22), and the convergence threshold is chosen to be  $5M\sigma^2$ , where  $M\sigma^2$  is the average array noise power. However, minor modifications have also been included to deal with special cases.

Table 1: Comparison of CRB and RMS errors obtained from 100 trials for SNR = 100 dB, with 8 sensors.

$$\theta_1 = 0^\circ, a_1 = 1\angle 0^\circ, \theta_2 = 10^\circ, a_2 = 0.5\angle 0^\circ.$$

$\beta$	$\theta_1(\text{deg.})$		$\theta_2(\text{deg.})$		Number of iterations
	CRB	RMS	CRB	RMS	
0.5	0.1064	0.5674	0.2127	1.2445	3-4
1.0	0.0263	0.0812	0.0527	0.1558	2
1.5	0.0115	0.0550	0.0231	0.1084	3
2.0	0.0064	0.0197	0.0127	0.0374	7
2.5	0.0040	0.0196	0.0079	0.0317	25-35
3.0	0.0027	0.0951	0.0054	0.1854	100

Table 2: Comparison of CRB and RMS errors obtained from 100 trials for SNR = 80 dB, with 8 sensors.

$$\theta_1 = 0^\circ, a_1 = 1\angle 0^\circ, \theta_2 = 10^\circ, a_2 = 0.5\angle 0^\circ.$$

$\beta$	$\theta_1(\text{deg.})$		$\theta_2(\text{deg.})$		Number of iterations
	CRB	RMS	CRB	RMS	
0.5	1.0635	—	2.1270	—	—
1.0	0.2634	0.8230	0.5269	2.0933	1-3
1.5	0.1153	0.2239	0.2307	0.4374	2-3
2.0	0.0636	0.1525	0.1271	0.2730	5-7
2.5	0.0397	0.1897	0.0793	0.2899	25-45
3.0	0.0267	0.2375	0.0535	0.4559	50-100

The "—" entries in Table 2 indicate that the corresponding results are not reliable, since the algorithm did not work satisfactorily for that case. Therefore, these results are not included.

The SNRs below 80 dB were not considered, since the CRBs for lower SNR become excessive. For ex-

ample, the CRB for  $\theta_2$  will be 5.27 for SNR = 60 dB and  $\beta = 1$ . This shows that at least 70-80 dB SNR is required for acceptable resolution in this specific problem. If the DOAs are more separated, operation down to about 40 dB is possible.

#### Observations on the results:

- The CRB decreases as  $\beta$  increases, approximately by  $1/\beta^2$ . This behavior is expected since a larger array dimension provides better resolution.
- For moderate or small values of  $\beta$ , the ratio of the RMS error to the CRB, for each DOA, is between 2 and 6. This deviation is explained by the fact that the CRBs are obtained by considering all the parameters at the same time, i.e., assuming a joint estimation. However, the proposed algorithm is based on parameter separation which is inherent in the LP method. Therefore, the estimates of DOAs and complex envelopes cannot be obtained simultaneously. The estimates of complex envelopes are only used to estimate the distortion and to improve the results. The reason why CRB is considered is that it provides a frame of reference, i.e., the best resolution theoretically achievable. However, all suboptimal (and practical) methods can only attain errors which can be considerably higher than CRB.
- The ratio of the RMS error to the CRB is generally smaller for SNR = 80 dB than that for SNR = 100 dB, for a given  $\beta$  and for each DOA. It is conjectured that for SNR = 100 dB the distortion term dominates, whereas for SNR = 80 dB the noise term is dominant.
- For large values of  $\beta$ , the ratio of the RMS error to the CRB, for each DOA, becomes larger. This is due to the increase in the distortion. Therefore, the algorithm in its current form seems to be suitable for the cases in which the array sizes should be relatively small such as in HF and VHF DF systems on mobile platforms.
- Although not shown in the tables, similar observations have been made for the estimates of complex envelopes.
- For small or moderate values of  $\beta$ , the algorithm converges in a few iterations. However, the average number of iterations required for convergence increases significantly as  $\beta$  increases. Actually, an oscillatory behavior is observed for such cases. For example, for  $\beta = 3$  most of the iterations were stopped without satisfying the aforementioned criterion since the maximum iteration limit was exceeded. This is thought to be due to the fact that the distortion term gets larger as  $\beta$  grows.



## 5. CONCLUSIONS

A new algorithm for estimating the directions of multiple emitters by employing a circular array which uses a single snapshot is proposed. The algorithm is inherently insensitive to correlation among the emitters, since it uses a single snapshot. The geometry is suitable for complete azimuthal coverage.

It is observed through computer simulations that the performance of the algorithm is quite satisfactory when the array circumference is in the order of the wavelength. Thus, the method seems to be suitable for, for example, mobile platforms in HF-VHF bands.

Currently, the problem of obtaining convergence for larger array sizes is being studied. Also, performance of the algorithm for sensor patterns other than the omnidirectional will be investigated.

## References

- [1] Ziskind, I. and Wax, M., "Maximum Likelihood Localization of Multiple Sources by Alternating Projection", *IEEE Trans. Acoust., Speech, Signal Processing*, **36**, 10, Oct. 1988, pp 1553-1560.
- [2] Feder, M. and Weinstein, E., "Parameter Estimation of Superimposed Signals Using The EM Algorithm", *IEEE Trans. Acoust., Speech, Signal Processing*, **36**, 4, Apr. 1988, pp 477-489.
- [3] Burg, J.P., "Maximum Entropy Spectral Analysis", Ph.D. dissertation, Stanford Univ., Stanford, CA, 1975.
- [4] Jaynes, E.T., "On The Rationale of Maximum Entropy Methods", *Proc. IEEE*, **70**, 9, Sept. 1982, pp 939-952.
- [5] Tufts, D.W. and Kumaresan, R., "Estimation of Frequencies of Multiple Sinusoids : Making Linear Prediction Perform Like Maximum Likelihood", *Proc. IEEE*, **70**, 9, Sept. 1982, pp 975-989.
- [6] Johnson, D.H., "The Application of Spectral Estimation Methods to Bearing Estimation Problems", *Proc. IEEE*, **70**, 9, Sept. 1982, pp 1018-1028.
- [7] Haykin, S., "Radar Array Processing for Angle of Arrival Estimation", in "Array Signal Processing", Haykin, S., Ed., Englewood Cliffs, NJ, Prentice-Hall, 1984, pp 194-292.
- [8] Schmidt, R.O., "Multiple Emitter Location and Signal Parameter Estimation", *IEEE Trans. Antennas Propagat.*, **34**, 3, Mar. 1986, pp 276-280.
- [9] Kumaresan, R. and Tufts, D.W., "Estimating The Angles of Arrival of Multiple Plane Waves", *IEEE Trans. Aerosp. Electron. Syst.*, **19**, 1, Jan. 1983, pp 134-139.
- [10] Reddi, S.S., "Multiple Source Location - A Digital Approach", *IEEE Trans. Aerosp. Electron. Syst.*, **15**, 1, Jan. 1979, pp 95-105.
- [11] Roy, R.H., "ESPRIT - Estimation of Signal Parameters via Rotational Invariance Techniques", Ph.D. dissertation, Stanford Univ., Stanford, CA, 1987.
- [12] Evans, J.E., Johnson, J.R., and Sun, D.F., "High Resolution Angular Spectrum Estimation Techniques for Terrain Scattering Analysis and Angle of Arrival Estimation", in *Proc. 1st ASSP Workshop Spectral Estimation*, Hamilton, Ont., Canada, 1981, pp 134-139.
- [13] Shan, T.J., Wax, M., and Kailath, T., "On Spatial Smoothing for Direction-of-Arrival Estimation of Coherent Signals", *IEEE Trans. Acoust., Speech, Signal Processing*, **33**, 4, Aug. 1985, pp 806-811.
- [14] Williams, R.T., Prasad, S., Mahalanabis, A.K., and Sibul, L.H., "An Improved Spatial Smoothing Technique for Bearing Estimation in a Multipath Environment", *IEEE Trans. Acoust., Speech, Signal Processing*, **36**, 4, Apr. 1988, pp 425-432.
- [15] Pillai, S.U. and Kwon, B.H., "Forward-Backward Spatial Smoothing Techniques for Coherent Signal Identification", *IEEE Trans. Acoust., Speech, Signal Processing*, **37**, 1, Jan. 1989, pp 8-15.
- [16] Cadzow, J.A., "A High Resolution Direction-of-Arrival Algorithm for Narrow-Band Coherent and Incoherent Sources", *IEEE Trans. Acoust., Speech, Signal Processing*, **36**, 7, July 1988, pp 965-979.
- [17] Kay, S.M., "Modern Spectral Estimation: Theory and Application", Englewood Cliffs, NJ, Prentice-Hall, 1988 (ISBN 0-13-598582-X).
- [18] Marple, S.L.Jr., "Digital Spectral Analysis with Applications", Englewood Cliffs, NJ, Prentice-Hall, 1987 (ISBN 0-13-214149-3).
- [19] Nikias, C.L. and Scott, P.D., "The Covariance Least-Squares Algorithm for Spectral Estimation of Processes of Short Data Length", *IEEE Trans. Geosci. Remote Sensing*, **21**, Apr. 1983, pp 180-190.
- [20] Van Trees, H.L., "Detection, Estimation, and Modulation Theory : Part I", NY, Wiley, 1968.

#### **DISCUSSION**

**G. MULTEDO**

This method is an extension of the linear prediction for linear arrays to circular arrays. Have you taken the effect of correlation in the multipaths into account?

**AUTHOR'S REPLY**

The method is insensitive to correlation in the paths.

**R. JENKINS**

At HF, the calibration of antennas can be a problem. Have you considered the effect of errors or uncertainties in the gains of the individual element antennas on your results?

**AUTHOR'S REPLY**

This problem has not been considered yet.

# Automatic Feature Extraction and Localization using Data Fusion of Radar and Infrared images

Stephane HOUZELLE  
AEROSPATIALE, Service E/ETRI  
1 Rue Pablo Picasso  
F-78114 Magny-les-Hameaux

Gerard GIRAUDON  
INRIA  
BP 93  
F-06902 Sophia Antipolis Cedex

## Abstract

*This paper presents two examples of low level strategies using multisensor data fusion, one for bridge extraction, and one for urban area extraction. These extractions are made from a couple of coregistered Synthetic Aperture Radar (SAR) and SPOT images. These features are very different by their dimensions, their shape, and their radiometry. Thus we can prove the reliability of our approach on various types of features. Our method uses the notion of complementarity of each sensor, and the notion of context in the observed scene. For bridge detection, we first segment water in the SPOT image, to spatially constrain the bridge research in the SAR image. This research is achieved using a correlation method. To detect an urban area, we first use the knowledge that it produces very bright texture in SAR imagery. Thus, the main part of urban backscatters is extracted using an adaptive thresholding which keeps the upper band of the gray level histogram of the SAR image. This mask is then used for classification as a training mask of urban area texture in SPOT image. We determine the non urban zone training set using a distance map of the urban training zone boundaries. Classification is performed with a multivariate Gaussian classifier. The results we obtained are very encouraging, especially if we consider the robustness of the bridge detection method.*

## 1 INTRODUCTION

For many years, remotely sensed images with high resolution (10-20m) have proven their usefulness for land description and scene analysis. Nevertheless, high resolution implies high complexity in images, making the analysis more difficult. Further more, high resolution forward the detection of objects which are more and more complex and/or small. For those kinds of objects, sensor description reliability can be very poor and low-level techniques, commonly employed to extract those objects, suffer from this restricted description.

So it becomes a necessity to use multisensor systems to have different points of view from the scene. In this way, you can use the best of each sensor to complete the deficiency of the others, or use the redundancy between the sensors to increase the reliability of each.

But to best use complementarity and redundancy of sensors, we need to know exactly the nature of each sensor, its geometry, and the way images are acquired and computed.

In this paper, we will focus on data fusion of coregistered Synthetic Aperture Radar (SAR) images with images obtained in the optical band or the near infrared with the French satellite SPOT.

Satellite SAR is an active system. It employs coherent microwaves to generate images of terrain. Each pixel of the image is the result of the estimation of the reflectivity distribution of the terrain at a given wavelength, polarization and viewing angle. Because SAR provides its own source to illuminate the target, it can be operated independently of weather influenced solar illumination, and hence has a great advantage over optical sensors. However, the usefulness of SAR data for automatic extraction is limited by geometric distortion, speckle noise due to coherent acquisition, number of effective grey levels available to represent terrain feature, and influence of illumination parameters (polarization, viewing angle...) over backscatters.

Our objective is to make a system of scene analysis with improved performances and reliability, compared to an architecture using a single sensor. In this paper, we focus on bridge detection, and urban area extraction. This two types of features are very different by their dimensions, their shape, or their radiometry, so we can prove the reliability of our approach on various types of features.

After a brief review of previous works in multisensor data fusion (section 2), we will describe in detail, in section 3, the two low-level strategies we chose to illustrate our fusion methodology. Then different results are presented in section 4 and a short presentation of our future works will end this article.

## 2 PREVIOUS WORK

Luo *et al* [14] make a distinction between two different types of studies in multisensor fusion. The first one, called *data fusion* refers to each step of any process that exploit multisensor data for a given purpose since this step combines informations coming from different sources. The second called *data integration* is defined as the entire structure that manages each data fusion step.

#### Data Fusion :

One part of multisensor or multispectral data exploitation has been made to improve manual human interpretation of scenes. Commonly used approaches merge data from different sensors using the specificity of each. For example, Chavez, in [1], merges spatial information of a sensor with spectral information of an other while Welch *et al* [17] use the Intensity Hue Saturation (I.H.S.) color transform described in [8] to extract cartographic feature from Shuttle Imaging Radar-B (SIR-B) images and Landsat Thematic Mapper images.

On the other hand, work on machine oriented data analysis using multisensor fusion can be divided into two parts. The first one deals with probabilistic approaches. One method, the simplest, consists in forming a vector of data per pixel, each component of the vector referring to one source. These vectors are then treated by a classifier as attribute vectors of a single source [18]. Lee *et al* [11] suggest a more general scheme employing a global membership function, (similar to a joint posterior probability) derived from every source. However, the problem of this approach is that it requires similarity among sources and probabilistic modeling which are not suitable, for instance, to mix spectral data and elevation data [11].

Thus, a second approach has been studied. This one, called the evidential approach, uses the Dempster-Shafer theory of evidence. This method takes advantage on probabilistic one by representing ignorance. Supporting evidence, plausible evidence, uncertainty of decision and conflict between sources are deduced from a mass function (that represents evidence) and Dempster rules. In [6], the formal method is introduced, and two computationally efficient approaches to the theory are presented in [11] and [12].

#### Data Integration :

In [15], an architecture using a whiteboard (blackboard which supports parallelism in the knowledge source modules) is used. In this architecture, each low level specialist processes naturally one sensor set of data. Then the results of those specialists are combined for interpretation. In general, guidance of the system is as much top-down as possible, and so, exploits complementarity of sensors rather than competitiveness between them. This architecture is tested in the context of a robot vehicle called NAVLAB.

In [7], the fusion of the information is made at a mid-level of interpretation. Features of different types called globally "tokens" are combined using a constraint-based approach. Application uses region and line tokens for texture measure.

At last, [10] presents a complete neural architecture which contains Autonomous Agent whose functions are to use explicit description of context to classify situations and objects, and to perform some action in response to it.

Methods we are now going to present are *Data integration* methods (as previously called). As in [7],

fusion of the information is made at a mid-level of interpretation. Nevertheless, extracted features are not combined, but are used to detect other features. Like this, we fully exploit the complementarity of our sensors.

### 3 EXAMPLES OF LOW-LEVEL STRATEGIES

Now we focus on two interesting examples of feature extraction using data fusion. The aim is to extract bridges and urban areas. These features are interesting for their spatial differences : a bridge is a microscopic object in the image while urban area is a macroscopic object. In our work, we assume that SAR images are exactly superimposed to SPOT images. So images are registered.

#### 3.1 Bridge extraction

The bridge detection problem is very hard to solve using only one sensor : in multispectral SPOT images, information about bridges is absent (in SPOT XS3) or very weak (in SPOT XS1 and XS2). In SAR images backscatters of bridges can be very high but shapely not very characteristic. Thus the multisensor approach becomes a necessity to make an effective and reliable detection. Moreover, bridge dimensions are near the resolution limit of the sensors we use to detect them, and we have to use the context in which we find bridges. This context is mainly represented by two kinds of features : roads and rivers. We are now going to present the bridge detection strategy in the context of rivers. Water happens to be very easy to detect in near infrared images in which it appears very dark, and we know that the most reliable sensor to detect bridges is the radar. So our method proceeds in two steps :

- First we detect the river in the SPOT XS3 image.
- Then we use this segmentation as a research mask in the SAR image to detect bridges.

The detection of bridges in SAR image is then spatially constrained and becomes easier. As we can see, the base of our method is the complementarity of the sensors we use.

##### 3.1.1 River segmentation

We use a region growing algorithm. The principle of region-growing is, from a seed point, to spread under constraint the initial region (pixel) to boundary pixels of the region (neighbours). The constraint we set, uses a modeling of the energy repartition in the grey-level cooccurrence matrices (joint histogram) of the region being grown [9]. Because we take pixels with darkest grey levels as seed points, this methods segments dark regions of the image.

Then we have a pattern recognition step to perform. To do that, we use the skeleton of objects previously detected. Assuming that rivers are quite long and thin, the recognition attribute is defined as :

$$A = \frac{\text{pixel length of the skeleton}}{\text{number of branches of the skeleton}}$$

For river, this ratio may be high while for compact objects, which have many small branches, it may be quite small.

The last step is a merging step. A river could have been partially detected, and be formed by more than one object. To group these objects, we use their skeleton again, checking if two ends of skeletons might not have the same direction, an opposite sense, and the same alignment. If this case, we join the two ends of skeletons with a line and pixels of this new line are considered as seed points for a new region growing process, spatially constrained inside the englobing box of the line.

We thus obtain a research mask for bridges in the SAR image. Results are presented in section 4.1 and on figure 3.

### 3.1.2 Bridge detection

For bridge detection, a skeleton of the research mask is very useful, because the skeleton's perpendicular gives an approximative position and direction of a possible bridge. If a bridge happens not to be perpendicular to the river, its backscatter shape, which widens at the bank of the river, may allow the skeleton's perpendicular to overlap the main part of the backscatter. An example of bridge gray level distribution in a radar image is presented on figure 1.

0	0	0	95	91	85	82	80	78	77
0	0	0	91	86	81	80	80	79	79
0	0	0	88	83	78	77	79	79	79
0	0	95	92	87	81	78	78	78	79
0	0	104	103	98	93	86	82	80	82
0	107	107	109	108	104	98	92	88	90
108	106	107	110	113	114	113	109	106	105
105	103	103	107	112	119	123	120	118	0
92	92	9	101	110	120	126	0	0	0
85	83	86	95	107	120	0	0	0	0
78	78	82	93	106	121	0	0	0	0
74	76	84	95	105	0	0	0	0	0
76	80	88	100	0	0	0	0	0	0
80	84	92	0	0	0	0	0	0	0
85	91	0	0	0	0	0	0	0	0

Figure 1: Example of spatial gray level distribution of a bridge in a SAR image. 0 value represents pixels out of the river mask.

However, the main problem encountered in the bridge detection is the backscatter dependence from the incidence angle of the radar beam, and from the

direction of the bridge with regard to the illumination direction. The result of this dependence is a possible disappearance of the bridge backscatter.

To manage this difficulty, we use a correlation method with the following correlation vector :

$$(-2 \ -1 \ 2 \ 2 \ 2 \ -1 \ -2)$$

which is a ridge detector adapted to bridge gray level section.

The method consists in computing the correlation in every pixel of the skeleton's perpendicular, and to sum all these results. At every pixel, the correlation vector is positioned in the skeleton direction. We thus obtain one measure per pixel of the skeleton (Fig 2).

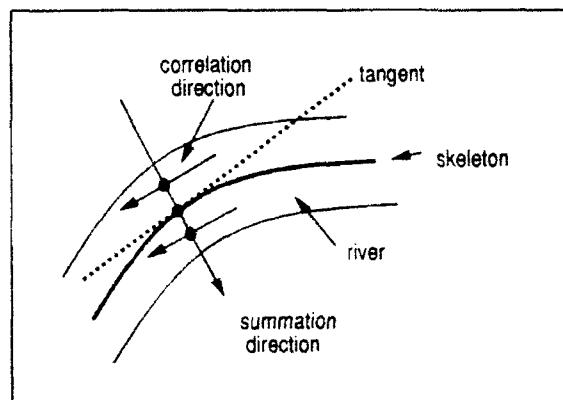


Figure 2: Correlation of bridge backscatters : the direction of the correlation vector at each point of the skeleton's perpendicular is shown. All correlation results on the perpendicular are summed.

To minimize the quantification problem of convolution direction, we locally rotate a subimage containing the neighborhood of the skeleton point, so that the tangent of the skeleton becomes parallel to the image rows. A planar interpolation is performed during this step. Like this, each skeleton's perpendicular becomes parallel to the rotated image columns, and correlation becomes easier to process.

The advantage of this method is to give a sort of bridge position probability and not to reject bridges which backscatter is partial. Results are presented on Figure 4 in section 4.1.

### 3.2 Urban Area detection

We set the urban area detection as a texture discrimination problem. Texture discrimination is a very difficult problem, and extensive research has been made in this field. Approaches are commonly characterized by two features : the texture measures, and the algorithm of classification. For supervised method, a third feature is added : the training zone determination. Supervised methods are generally more robust since we introduce in the process our knowledge of

the scene. Nevertheless, training zone determination make the method becoming semi-automatic.

The main contribution of multisensor data fusion in supervised methods is to automate the training zone determination.

In the most general form of supervised classification, training set determination has to be made so that :

- Information from one training zone is characteristic of one class,
- Information from each zone follows the statistical modeling you made (if any),
- All classes are represented in the training set.

However, the problem addressed here is slightly different from the general one, since only one class is important : the urban class. Thus, the first recommendation is not very important except for the urban class. We are now going to see how our method takes into account the other recommendations to determine automatically training zones for classification. We use a multivariate Gaussian Bayesian classification. However, any other supervised method using a statistic modeling of attributes distribution could replace our Bayesian classification process.

We know that urban areas produce very brilliant texture in SAR imagery. So, assuming that gray level distribution of the SAR image has a mean  $m$  and standard deviation  $\sigma$ , we extract the main part of urban backscatter by thresholding the image with a threshold value  $m + k * \sigma$ , where  $k$  is a constant. In the Gaussian assumption, this threshold gives a probability of rejection (false alarm) which is independent of  $m$  and  $\sigma$ , but only  $k$  dependent. After removing small regions and filling small holes of remaining regions, we obtain a partial mask of the urban area which is used as a training mask of urban area texture in a SPOT image.

However, some problems still remain to determine the non-urban zone training set. Three solutions can be considered : the first one consists in taking the darkest part of the SAR image and making the assumption that these pixels are not urban area type. Zones defined in this way are very easy to obtain, by thresholding the image as previously, with a threshold value set to  $m - k * \sigma$ . Nevertheless, it is then very difficult to check the reliability of the textural information contained in these zones.

The second method consists in considering only one texture type (urban) but introducing a reject threshold during the classification. This method uses only our knowledge of the scene but requires the choice of the reject threshold value, which is scene dependent.

The third method (that we have kept) consists in using a distance map of the urban training zone boundaries. This map gives, for each pixel outside urban training zones, the minimum distance to a boundary of one of these zones. Local maxima of this map are the

further pixels of potential urban areas, and thus represent the most probable sites of non-urban areas. Thus, we consider these local maxima as centers of training zones of non-urban areas. Every pixel of these zones are then determined assuming they are circular with a given surface value. The advantage of this method over the first one is that we control the number of zones and their surfaces. So, we can easily check their reliability for classification.

As we have already said, classification is achieved by using a multivariate Gaussian Bayesian classifier. The criterion we use to evaluate the reliability of non-urban training zones is the likelihood of each attribute distribution to the Gaussian model. A zone is kept for training only if all attributes distribution follow a Gaussian model.

So we make the training step of the classification totally automatic. Thus we can use a robust but semi-automatic supervised method as an automatic unsupervised method.

After experiments, we have selected four attributes for classification. The first two were introduced by Lowitz [13], and use local histograms of an image. They are called state (or phase), and module of a local histogram. The state is the dominant radiometry index of the local histogram, while the module is the sum of the square differences between normalized histogram values and flat histogram value. Let  $N$  be the surface of a local window  $W$ , and  $h_i$  the normalized histogram value for radiometry index  $i$  :

$$mod(h_W) = \sum_i \left( \frac{h_i}{N} - \frac{1}{N} \right)^2$$

The third attribute we use for classification is a Laplacian measure, and the fourth is a gradient orientation measure. This last is calculated using the mean gradient vector on a local window  $G_{mod}$ , and is defined as :

$$Grad\_Orient_W = E[\|G - G_{mod}\|^2]$$

## 4 RESULTS

The coherence needed to produce SAR images introduces a speckle noise increasing the difficulty of image processing. In all our studies we have filtered the SAR images with the Frost algorithm [3].

### 4.1 Bridge detection

Method presented in section 3.1 has been tested with a couple SPOT and SAR SIR-B images representing a site of a region in the east of France. A river (the Rhin) appears in the middle of images. Figure 3 shows the entire process that leads to the river mask : picture (a) shows the original SPOT XS3 image whose resolution is 20 m. Picture (b) shows the result of the dark region segmentation. The only parameter of the segmentation algorithm is a smoothing parameter,

totally independent from the scene. A morphological opening is then processed (picture (c)) to smooth the result, and skeleton is calculated (picture (d)). We use this skeleton to recognize the river. The pattern recognition parameter  $A$  (section 3.1) set to 50, is enough to perform this step. Then we try to connect the ends of the remaining skeletons. The result is showed on picture (e). The final result after the new region growing pass is presented on picture (f).

The result of bridge detection in the SAR image is presented on figure 4. Picture (a) represents the original image with circles around bridges. We can see the problem of backscatter disappearance on the bridge located in the near center of the image. (b) shows the research mask, and (c) represents the result of correlation. During correlation, the value of pixels out of the research mask was fixed to the mean of the research mask. As we can see, detection of the problematic bridges is possible with this result (d), and, all the bridges have been detected.

We have tested this method on a coast image where wind makes river gray levels very different, and, again, all bridges were detected without any false alarm, proving the robustness of the method.

## 4.2 Urban Area detection

To test our method described in section 3.2, we have used a couple of SPOT SAR Seasat images containing a coast town. Figure 5.(a) shows the original SAR image which is very noisy though it was speckle filtered. (b) shows the threshold result with  $k = 1$ . Picture (c) represents the final training set we have used for classification. circular regions are non urban area training zones. At this step, we have suppressed small regions of urban training set (whose maximum distance to boundaries inside the region inferior to 6 pixels) and fill up holes (which surfaces were inferior to 70 pixels). Picture (d), shows the original SPOT XS3 image used for classification, and picture (e) shows a urban area mask, determined manually on a cartographic map of the region. Finally, picture (f) shows the final result after classification and cleaning (as performed previously for primary urban mask).

## 5 CONCLUSION

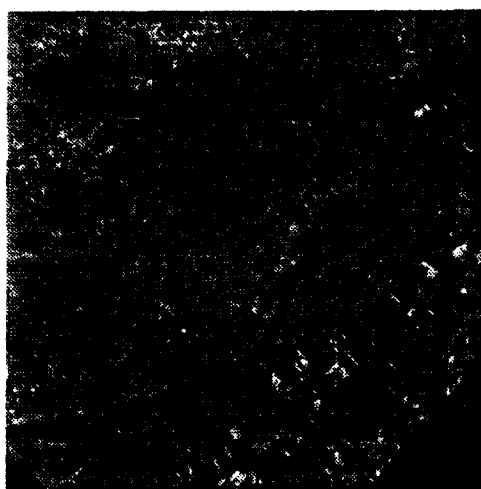
This paper presented two low level strategies we have used for bridge and urban area extraction. These methods take explicitly into account the complementarity of sensors and our knowledge about object gray level distribution in each of the sensors. The results we obtained are very encouraging and further works have to be done to integrate these low level strategies in a general system of scene analysis. The architecture of such a system is fundamental, and has to managed the communication requests of all part of the system, while being adaptable enough to allow integration of new entity. That is why we choose a Blackboard ar-

chitecture for our system. This architecture uses the concepts of the system called MESSIE (Multi Specialist Expert System for Scene Interpretation and Evaluation), which was successfully tested on roads and buildings detection in aerial images [4], [5]. Such systems are adapted to multisensor data fusion in the sense that they allow an easy incorporation of new knowledges, and an easy cooperation between specialists. The global strategy of our system is to use the context of the objects we want to extract, and the idea of evident objects. These two features help the system to find, at any moment of the reasoning, its potentiality to detect an object.

## Acknowledgments

The authors would like to acknowledge the contributions of Valerie Gertner and Herve Andre in the choice of urban area texture measures, and Philippe Nonin (ISTAR company) in the classification process.

- [1] P. CHAVEZ Jr. *Digital Merging of Landsat TM and digitized NHAP data for 1:24,000-scale image mapping*, Photogrammetric Engineering and remote sensing, Vol 52, No 10, 1986, pp 1637-1646.
- [2] R. DERICHE, *Optimal edge detection using recursive filtering*, Proc. of the first ICCV, London, June 1987.
- [3] V. FROST, J. ABBOTT STILES, K. SHANMUGAN, J. HOLTZMAN, *A model for radar images and its application to adaptive digital filtering of multiplicative noise*, IEEE Trans. on PAMI, Vol PAMI-4 No 2, March 1982.
- [4] P. GARNESSON, G. GIRAUDON, P. MONTESSINOS, *Detecting Buildings and roads in aerial images by a multi expert system*, Les systemes experts et leurs applications, France, Avignon 1989.
- [5] P. GARNESSON, G. GIRAUDON, P. MONTESSINOS, *An image analysis system, application for aerial imagery interpretation*, Int. Conf. on Pattern Recognition, 1990.
- [6] T. GARVEY, J. LOWRANCE, M. FISCHLER, *An inference technique for integrating knowledge from disparate sources*, Proc. of the 7th Int. Joint Conf. on Artificial Intelligence, Vancouver 1981, pp 319-325.
- [7] R. HANSON, E. RISEMAN, T. WILLIAMS, *Sensor and information fusion from knowledge-based constraints*, Proc. of SPIE Vol 931 Sensor Fusion 1988, pp 186-196.
- [8] R. HAYDN, G. DALKE, J. HENKEL, *Application to the IHS color transform to the processing*



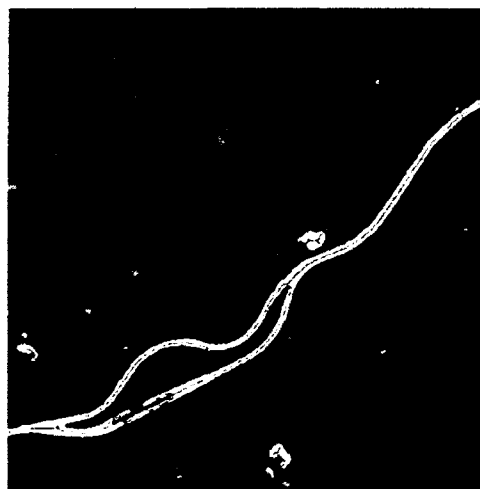
(a)



(b)



(c)



(d)



(e)



(f)

Figure 3. Complete processus that lead to river mask. (a) original SPOT XS3 image. (b) Result of the region growing. (c) smoothing of the result. (d) object's skeletons. (e) remaining skeleton after recognition and prolongation. (f) final mask



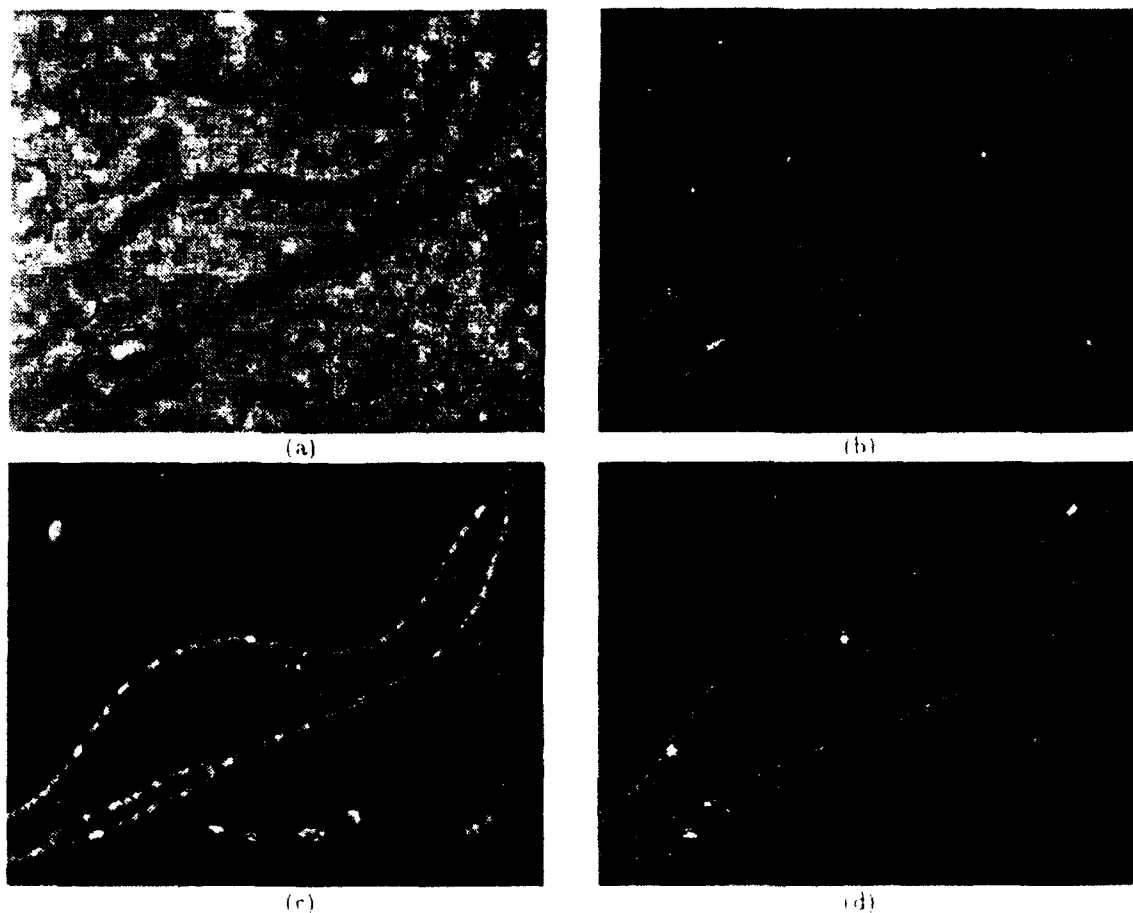


Figure 4: Detection of bridges in SAR image : (a) Original image. Circles show bridge positions (b) research mask, (c) result of the correlation, (d) final result, pic detection.

- of multisensor data and image enhancement*, Int. Symp. on Remote sensing of Arid and Semi-Arid Lands, Cairo, 1982, pp 599-616.
- [9] S. HOUZELLE, G. GIRAUDON, *Model based region segmentation*, Workshop on computer vision and image processing for space-born applications, Noordwijk, June 91.
- [10] O. JAKUBOWICZ, *Autonomous reconfiguration of sensor systems using neural nets*, Proc. of SPIE Vol 931 Sensor Fusion 1988, pp 197-203.
- [11] T. LEE, J. RICHARDS, P. SWAIN, *Probabilistic and evidential approaches for multisource data analysis*, IEEE Trans. on Geoscience and remote sensing, Vol GE 25, No 3 1987, pp 283-293.
- [12] N. LEHRER, G. REYNOLDS, J. GRIFFITH, *Initial hypothesis formation in image understanding using an automatically generated knowledge base*, Workshop on Image Understanding 1987, pp 521-537.
- [13] G. LOWITZ, *Mapping the local information content of a spatial image*, Pattern Recognition, Vol 17, No 5, 1984, pp 545-550.
- [14] R. LUO, M. KAY, *Multisensor integration and fusion: issues and approaches*, Proc. of SPIE Vol 931 Sensor Fusion 1988, pp 42-49.
- [15] S. SHAFER, A. STENZEL, C. THORPE, *An architecture for sensor fusion in a mobile robot*, Int. Conf. on Robotics and Automation, San Francisco 1986, pp 2202-2011.
- [16] P. SWAIN, *Multisource data analysis in remote sensing and geographic information processing*, 1985 Machine Processing of Remotely Sensed Data Symposium, pp 211-218.
- [17] R. WELCH, M. EHLERS, *Cartographic feature extraction with integrated SIR-B and Landsat TM Images*, Int. Journ. of Remote sensing, Vol 9, No 5 1988, pp 873-889.
- [18] S. WU, *Analysis of data acquired by shuttle imaging radar SIR-A and Landsat Thematic Mapper over Baldwin county, Alabama 1985* Machine Processing of Remotely Sensed Data Symposium, pp 473-482.

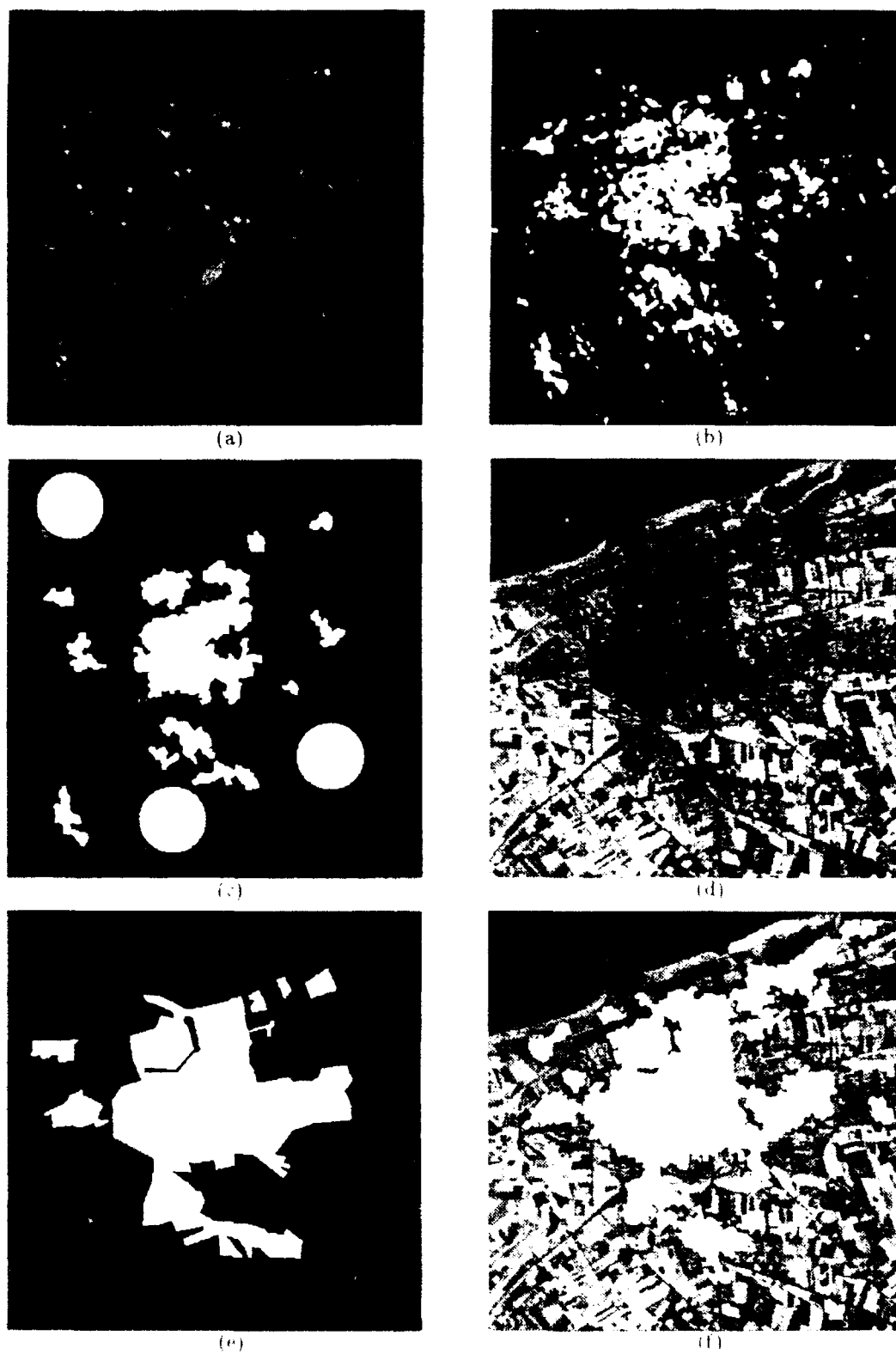


Figure 5. Urban area segmentation. (a) Original SAR Seasat image. (b) threshold result with  $k = 1$ . (c) cleaned mask of urban area, and non urban zone chosen for classification (circles). (d) Original SPOT XS3 image. (e) manually segmented mask of urban area. (f) final result after classification and cleaning.

## DISCUSSION

### E. SCHWEICHER

Pourriez-vous évaluer l'augmentation de la probabilité de détection et/ou la diminution de la probabilité de fausse alarme de la fusion de données par rapport aux probabilités correspondantes caractérisant le radar seul?

*Could you assess the increase in probability of detection and/or decrease in the probability of false alarm resulting from data fusion compared to the corresponding probabilities characterising the radar only?*

### AUTHOR'S REPLY

Stéphane Houzelle pourrait quantitativement répondre à la question. Qualitativement, l'amélioration est très importante quant à la robustesse de l'extraction en permettant de s'affranchir des problèmes dus à un simple seuillage.

*Stephane Houzelle could give a quantitative reply to this question. Qualitatively, the improvement is considerable with respect to the strength of the extraction, thereby overcoming any problems caused by simple thresholding.*

## DIGITAL PROCESSING FOR POSITIONING WITH ONE SATELLITE

by

A. Marguinaud  
Ground Systems Department  
Alcatel Espace  
11 Avenue Dubonnet  
92407 Courbevoie  
France

### SUMMARY

This paper is a product of theoretical and practical work done with geostationary and low earth-orbiting (SARSAT) satellites. Reliability, efficiency and cost considerations have resulted in an all digital implementation with standard processors.

The analog signal is represented by complex samples calculated from real sampling with an algorithm based on an original concept.

Any telecommunication (*or radar*) signal is represented by reversible transformations of a sinusoidal carrier. Maximum likelihood synchronization and demodulation must be used and consists of reconstructing the "best" carrier from the received signal. Depending on the uncertainty level, one tries a certain number of time-frequency hypotheses, and for each hypothesis, one makes optimum estimations of 3 carrier parameters : modulus, phase and frequency.

Since one deals with one circular function, it is straightforward to develop the phase, so that phase and frequency are the 2 parameters defining a straight line in the developed phase-time plane. The 2 corresponding estimators are decorrelated.

For a terrestrial vehicle, one can have a good estimation of its velocity, and even for a stationary satellite there is an apparent periodical known motion, which can be used to locate the vehicle by a convenient combination of elementary estimations. Of course some frequency spreading, clock stability and delay are necessary, but explicit performance is easy to derive as functions of operational conditions. Analysis of numerical results may suggest useful practical systems.

### I. INTRODUCTION

The SARSAT system has been designed to locate terrestrial radio beacons emissions (406 MHz) using the doppler measured by a low orbiting satellite (800 km).

For operational purposes (*alarm delay, beacon oscillator frequency*), however, it is useful to detect these signals in a stationary satellite and then transmit them to an earth station.

In this case, due to the greater distance, the received power level is considerably lower, so that it is necessary to be able to detect and demodulate separate messages (about 100 bits, biphasic L) of unknown phase carrier at an energetic signal to noise ratio  $E_b/N_0$  of about 0 dB.

Instead of improving established methods using phase lock loops for carrier and clock recovery, we use sampled analytical signal representation, maximum likelihood principles, statistical hypothesis and testing estimation to develop a new approach for message recovery which compares favourably with traditional methods [1].

These new methods have been implemented on a multiprocessor BULL SPS7 equipped with Motorola 68 000 and Texas TMS and tested with an operational geostationary satellite [2].

The modelling of the same estimation methods with a signal received by a low orbiting satellite has been completed with success and it has been verified that a negligible performance degradation is induced by the additional smooth doppler.

After a presentation of these methods and of their theoretical performances, we apply them to terrestrial location with a (*quasi*) stationary satellite.

## II. COMPLEX SAMPLING OF REAL SIGNALS

Complex representation of real signals has been introduced by Ville in 1948 [3] to give a permissible mathematical definition of phase and frequency for real one dimensional signals. This mathematical approach has proven to be very useful in digital signal processing implementation, because of its two equivalent representations of complex samples: cartesian and polar. Polar representation allows a simple algorithm to be used for changing carrier frequencies, and avoids dephasing. On the other hand, cartesian representation allows an efficient means (complex linear summation) to raise the signal to noise power ratio of received modulated signals. And lastly, going from one representation to the other requires few elementary operations thanks to precalculated tables.

The above considerations show the interest in having an efficient way of obtaining the complex representation of the real signal. First implementations were based on the modulation of a real signal by the two quadratures components of a signal delivered by an oscillator. The complexity of this method is due to the existence of a separate chain for each cartesian component. A digital implementation is preferable since both components can be generated from the same series of samples.

Instead of using the discrete Fourier transform or Hilbert transform, which present some problems of absolute convergence, we have developed a direct method for generation of complex signals which rests on a uniform representation of any continuous signal as a sum of trigonometric functions (Weistrass Theorem) [4].

Thanks to the above property one can approximate the real signal  $x(t)$  and its associated quadrature signal  $y(t)$  by (1).

(1)

$$x(t) = \sum_{k=0}^{\infty} a_k \cos(k\omega t + \phi_k) \Rightarrow y(t) = \sum_{k=0}^{\infty} a_k \sin(k\omega t + \phi_k)$$

This expression remains true if one considers any periodic time sampling such that  $t = m\Delta t$  ( $m$  integer),  $k\omega$  can be replaced by :

$$k\omega + \frac{2\pi}{\Delta t}k$$

for any common integer  $k$ , so that all terms of (1) can be interpreted as sinewaves belonging to the same frequency band :

$$\left( F - \frac{1}{2\Delta t}, F + \frac{1}{2\Delta t} \right)$$

The 2 complex components at the same time are calculated by 2 FIR filters (*finite impulse response*) :  $x(t)$  by a finite symmetric weighting  $y$ ,  $y(t)$  by a finite time antisymmetric weighting.

It is sufficient to optimise these 2 FIR for sines whose frequencies belong to the spectrum of received signals. We have derived an analytic theory which gives the number of coefficients and their values as a function of the specified quadrature noise to signal power ratio [5].

However we give here a direct derivation of the 4 FIR coefficients which are used to calculate the 2 quadrature components in the middle of 2 consecutive samples.

It will be clear later on that the "interpolating" analytic front-end has the advantage of allowing an efficient correction of the ADC offset.

Complexity analysis shows that the simplest design is obtained when  $F = \frac{1}{4\Delta t}$ .

This means that for a sine wave at the band centre, the signal vector turns by an angle equal to  $\pi/2$  between 2 consecutive samples.

To obtain the 4 coefficients, one has just to take 4 successive samples of  $\cos \frac{\pi t}{2\Delta t}$  and  $\sin \frac{\pi t}{2\Delta t}$  weighted by a semi period of  $\cos \alpha t$  to get a bandpass filter effect.

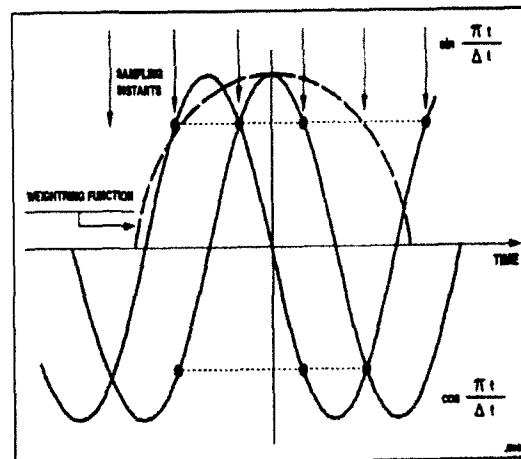


Figure 1 : Heuristic analytical head concept

By taking :

$$\alpha = \frac{2 \text{Arc cos} \sqrt{\frac{\pi}{6}}}{\Delta t}$$

one gets the following simple formulae for interpolating  $x(t)$  and  $y(t)$  from the real signal  $r(t)$ .

(2)

$$\begin{cases} x(t) = 3 \left[ r\left(t + \frac{\Delta t}{2}\right) + r\left(t - \frac{\Delta t}{2}\right) \right] - \left[ r\left(t + \frac{3\Delta t}{2}\right) + r\left(t - \frac{3\Delta t}{2}\right) \right] \\ y(t) = 3 \left[ r\left(t + \frac{\Delta t}{2}\right) - r\left(t - \frac{\Delta t}{2}\right) \right] + \left[ r\left(t + \frac{3\Delta t}{2}\right) - r\left(t - \frac{3\Delta t}{2}\right) \right] \end{cases}$$

To evaluate the quadrature defined by (2), it is sufficient to test it with the signal  $r(t) = \cos(\omega t)$ , recalling that  $\cos 3\theta = -3\cos\theta + 4\cos^3\theta$  and  $\sin 3\theta = 3\sin\theta - 4\sin^3\theta$ .

(3)

$$\begin{cases} r(t) = \cos \omega t \\ \theta = \frac{\omega \Delta t}{2} \end{cases} \Rightarrow \begin{cases} x(t) = 4(3\cos\theta - 2\cos^3\theta)\cos \omega t \\ y(t) = 4(3\sin\theta - 2\sin^3\theta)\sin \omega t \end{cases}$$

The factors of  $\cos \omega t$  and  $\sin \omega t$  are equal when  $\theta = \frac{\omega_0 \Delta t}{2} = \frac{\pi}{4}$ , which suggest introducing  $\Delta\omega$  such that :

$$\omega = \omega_0 + \Delta\omega \quad \text{and} \quad \theta = \frac{\omega \Delta t}{2} = \frac{\pi}{4} \left( 1 + \frac{\Delta\omega}{\omega_0} \right)$$

(4)

$$\begin{cases} x(t) = 4\sqrt{2} \left( 1 - \sin^2 \frac{\pi \Delta\omega}{2 \omega_0} \right)^{1/2} \left( 1 + \frac{1}{2} \sin^2 \frac{\pi \Delta\omega}{2 \omega_0} \right) \cos \omega t = X \cos \omega t \\ y(t) = 4\sqrt{2} \left( 1 + \sin^2 \frac{\pi \Delta\omega}{2 \omega_0} \right)^{1/2} \left( 1 - \frac{1}{2} \sin^2 \frac{\pi \Delta\omega}{2 \omega_0} \right) \sin \omega t = Y \sin \omega t \end{cases}$$

$X$  and  $Y$  are considered as the sum of the correct signal plus the "quadratic noise".

(5)

$$\begin{cases} (X^2 + Y^2) = 16 \left( 4 - 3 \sin^2 \frac{\pi \Delta\omega}{2 \omega_0} \right) = 16 \left( 1 + 3 \cos^2 \frac{\pi \Delta\omega}{2 \omega_0} \right) \\ (Y^2 - X^2) = 16 \sin^3 \frac{\pi \Delta\omega}{2 \omega_0} \end{cases}$$

The ratio  $\rho$  of the signal divided by the quadrature noise is deduced from (5) according to (6).

(6)

$$\begin{cases} X = G - \epsilon \\ Y = G + \epsilon \end{cases} \Rightarrow \begin{cases} X^2 + Y^2 = G^2 + \epsilon^2 \\ Y^2 - X^2 = 2G\epsilon \end{cases}$$

$$\rho = \frac{G^2}{\epsilon^2} = 4 \left( \frac{X^2 + Y^2}{Y^2 - X^2} \right) = 4 \left( \frac{1 + 3 \cos^2 \frac{\pi \Delta\omega}{2 \omega_0}}{\sin^3 \frac{\pi \Delta\omega}{2 \omega_0}} \right)^2$$

One verifies numerically that usable bandwidth is of the order of  $\omega_0$ .

Numerical application of (6) gives the signal over noise ratio in dB as a function of  $\frac{\Delta\omega}{\omega_0}$ , recall that the usable normalized bandwidth is  $\frac{2\Delta\omega}{\omega_0}$ , for a normalised real sampling frequency 4.

$\rho(\text{dB})$	10	20	30	40	50	60
$\frac{\Delta\omega}{\omega_0}$	0,78	0,55	0,39	0,27	0,18	0,12

Another interesting aspect of this analytical front-end, is that it improves the dynamics of the received signal, defined as the ratio of the expected square of the signal divided by the noise variance.

The real signal  $\cos \omega t$  has a power  $\frac{a^2}{2}$  and a noise characterised by its variance  $\sigma^2$ , but after the analytic quadrature stage, the signal power is :

$$16 \left( 4 - \sin^2 \frac{\pi \Delta\omega}{2 \omega_0} \right) \frac{a^2}{2}$$

and the noise variance is  $16\sigma^2$ , giving the dynamics improvement of  $\delta$  when the noise samples are independent.

(7)

$$\delta = 4 - 3 \sin^2 \frac{\pi \Delta\omega}{2 \omega_0}$$

### III. CARRIER FORMALISM OF MODULATED SIGNAL

The advantage of this representation is that synchronisation, demodulation and parameter estimation are reduced to estimations of the 3 characteristics of a sinewave. For instance, the detection of an awaited signal is decided when the estimated carrier amplitude exceeds a prescribed threshold. For position location, phase will be associated with range and frequency with a geometrical angle as we shall explain later.

Since we are dealing with sinewaves, operations such as phase development and parameter estimation can be implemented with a simple optimal algorithm. To justify these claims, we shall develop (*virtual*) carrier formalism to describe current modulation and spread spectrum.

The first example considered is FSK (frequency shift keying) in which one frequency sine wave among  $m=2^b$  is emitted during an elementary symbol representative of  $b$  bits. During the transmission of one symbol, the diagram of the developed phase as a time function is a straight line, the slope of which can assume  $m$  values.

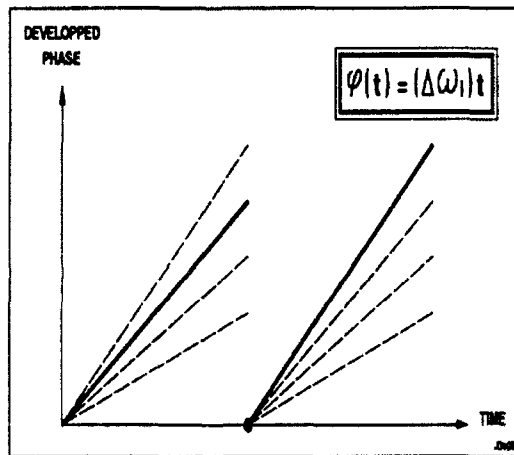


Figure 2 : FSK developed phase

One way to demodulate this signal when the symbol clock is known consists of applying the inverse phase rotation corresponding to the complex samples of the received signal and adding them during the symbol duration. The demodulated symbol is the one which gives the greatest modulus.

In practice, things are slightly different because the centre of the frequency band is different from zero and the symbol clock is unknown :

$$\phi(t) = \phi_0 + [\omega_0 + (\Delta\omega_1)](t - t_0)$$

The unknowns are  $\phi_0$ ,  $\omega_0$  and  $t_0$ . Generally, due to propagation hazards and modulation concept the value of  $\phi_0$  is not accessible, so that only  $\omega_0$  and  $t_0$  have to be evaluated. This evaluation is carried in 2 steps : hypothesis testing and followed by parameter estimation. Only hypothesis testing is considered in this paragraph, and it consists of calculating the carrier modulus for all starting time-carrier frequency couples corresponding to the estimated uncertainty and accepted performance degradation.

Assuming frequency spacing  $\Delta f$  and a time spacing  $\Delta T$  for a symbol duration  $T$ , the reduction factor of the carrier modulus evaluation is given by (8) :

$$(8) \quad R \leq \left(1 - \frac{\Delta T}{T}\right) \cos 2\pi(\Delta f)T$$

The next example concerns binary transmission by BPSK-PN spread spectrum (*binary phase shift keying pseudo noise*). The developed phase of this signal is represented on figure 3.

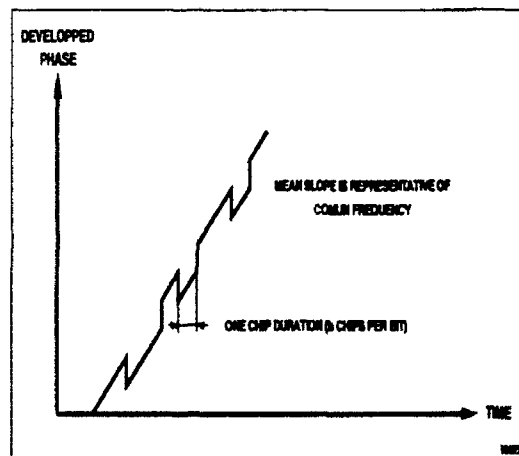


Figure 3 : BPSK - PN developed phase

The received signal undergoes adapted filtering before the complex sampling, the rate of which is such that there are 3 samples during each chip duration. Optimal processing requires the selection of one sample sequence from 3 for each frequency hypothesis which is obtained by selecting the sequence having the biggest estimated modulus.

A third example is MSK data transmission. As before one has to test a certain number of frequency hypotheses to have a sufficiently accurate estimate of the carrier amplitude (cf. (8)). The developed phase diagram is represented on figure (4).

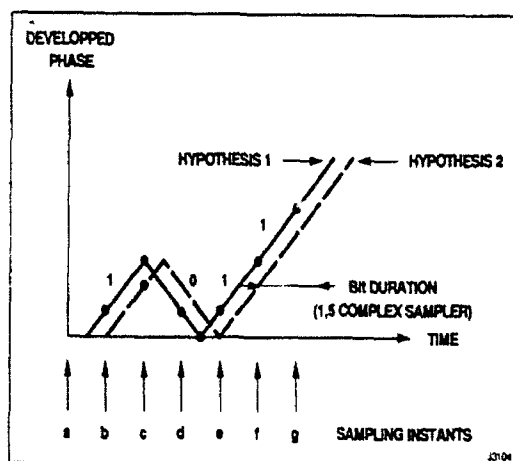


figure 4 : MSK developed phase

In order to test the sequence 1, 0, 1, 1 at each sampled time one tests 2 time phasings by applying the following rotations.

	b	c	d	e	f	g
I	$-\frac{\pi}{6}$	$-\frac{\pi}{2}$	$1 \frac{\pi}{6}$	$-\frac{\pi}{6}$	$-\frac{\pi}{2}$	$-\frac{5\pi}{6}$
II	0	$-\frac{\pi}{3}$	$-\frac{\pi}{3}$	0	$-\frac{\pi}{3}$	$-\frac{2\pi}{3}$

The sum of the 6 rotated vectors gives an estimate of the carrier sample for the sequence 1, 0, 1, 1. One continues these estimations for the  $2^4 = 16$  sequences, the greatest modulus indicates the transmitted 4 bits (1, 0, 1, 1 on fig.4) sequence. After synchronisation, it is possible to adjust the elementary rotations to the estimated clock phasing.

It is straightforward to extend the above procedure to combined phase and amplitude combined modulation. For all modulations schemes, we have developed algorithms to save a lot of elementary calculations (*frequency hypothesis, rotations, sums*) by re-using intermediate results so that implementation becomes quite acceptable.

The preceding examples have been given to show that any modulation or spread spectrum scheme can be considered as a succession of reversible complex transformations of a sine wave carrier.

#### IV. CIRCULAR PARAMETERS ESTIMATION PERFORMANCE

In the previous paragraph we have shown how to use a statistical hypothesis procedure to transform the received signal into a circular function perturbed by additive noise. For each hypothesis to be tested, the signal is represented by complex samples regularly spaced in time. For our purpose the problem is to estimate the 3 characteristics  $A$ ,  $\phi$ ,  $\omega$  in expressions of the form  $a \exp(j\omega t + \phi_0)$  perturbed by additive complex noise  $n = n_1 + jn_2$  verifying  $\langle n_1 \rangle = \langle n_2 \rangle = \langle n_1 n_2 \rangle = 0$  and  $\langle n_1^2 \rangle = \langle n_2^2 \rangle = \sigma^2$ .

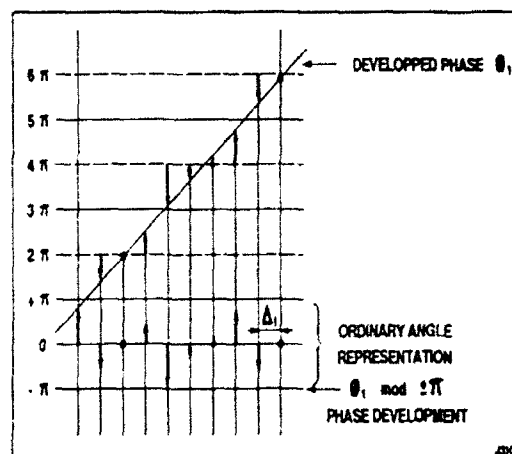


figure 5 : Carrier Phase development

The complex samples are of the form :

$A \exp(j\phi) + n_1 + jn_2$  with  $A$  and  $\phi$  unknown. The elementary estimation process replaces  $A$  by :  $A e^{\epsilon}$  and  $\phi$  by  $\phi + \Delta\phi$ .  
 $A \exp[\epsilon + j(\phi + \Delta\phi)] = A \exp j\phi + n_1 + jn_2$

By taking the module of both members, one obtains the next relation.

$$\frac{n_1^2 + n_2^2}{A^2} = e^{2\epsilon} - 2e^{\epsilon} \cos \Delta\phi + 1 \geq \sin^2 \Delta\phi$$

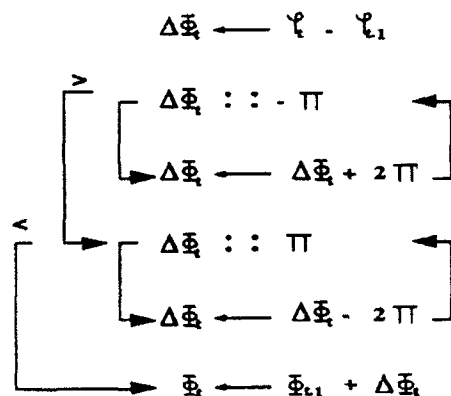
Averaging over noise level, one derives an upper bound for the variance of  $\sin \Delta\phi$ , which justifies relation (9) for practical evaluations.

$$(9) \quad \sigma_{\phi}^2 = \frac{2\sigma^2}{A^2} = \frac{1}{\rho}$$



In order to evaluate  $\omega$ ,  $\phi_0$ , and  $A$  characteristic of a carrier, one starts by constructing the **developed phase**  $\phi_1, \phi_2, \phi_3$  etc. According to algorithm (10) which connects the developed angles by the shortest path.

The samples are labelled by  $(r)$  considered as an integer.



: : is for comparison between left and right members

2199

The above algorithm fails if the angle jump between 2 consecutive samples has an amplitude greater than an absolute value of  $\pi$ .

Failure probability is an increasing function of the carrier rotation angle  $|\omega\Delta t|$  between 2 samples separated by  $\Delta t$ . To reduce this probability to its minimum, one subtracts a rotation  $\omega'$  from the unknown pulsation  $\omega$  such that  $|\omega - \omega'| \Delta t \ll 1$ .

Thus one needs some method which yields a reasonable approximation for the above frequency  $\omega$ . For this purpose one uses as before a frequency hypothesis technique and the optimisation of its implementation gives a Sliding Frequency Evaluator (SFE).

Given 2 successive samples  $\exp j(\omega t + \phi)$  and  $\exp j[\omega(t + \Delta t) + \phi]$  of the same carrier, we rotate onward the first by an angle  $\omega'\Delta t$  close to the actual (but unknown)  $\omega\Delta t$ . The "vectorial" sum of these 2 samples is interpreted as a sample at time  $t + \Delta t$  with an improved signal over noise ratio by a factor 2.

(11)

$$\begin{cases} \exp j(\omega t + \phi) + \exp j[\omega(t + \Delta t) + \phi - \omega'\Delta t] = K \exp j(\omega t + \phi) \\ K = 1 + \exp j(\omega - \omega')\Delta t = 2 + j(\omega - \omega')\Delta t \end{cases}$$

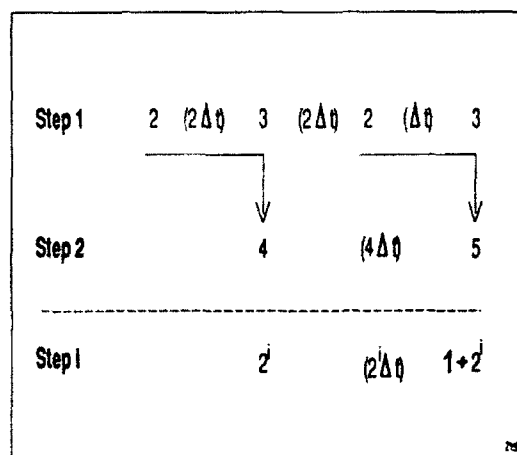
The main discrepancy with the expected factor 2 is the angle  $\frac{(\omega - \omega')\Delta t}{2}$ , it is important to note the sign of  $(\omega - \omega')$ .

For the first 2 samples we can take a number of regularly spaced frequencies determined by the frequency range uncertainty  $(2\Delta\omega)$ . For the next 2 samples, one does the same weighted sum but with a frequency hypothesis of different parity, in order to introduce some dithering in the frequency sampling. Usually, the frequency hypotheses will alternate from 2 to 3 :

$$\left(-\frac{\Delta\omega}{2}, \frac{\Delta\omega}{2}\right) \text{ and } \left(-\frac{2}{3}\Delta\omega, 0, \frac{2}{3}\Delta\omega\right)$$

At the second step, one takes 2 consecutive hypothesis sets (of 2 and 3 elements respectively) to calculate alternately consecutive sets of 4 and 5 hypotheses being separated by the time interval  $2\Delta t$ . At each new step the number of hypotheses has to be doubled, since the distance between 2 consecutive sets is double in order to keep the same angle uncertainty.

The evolution of the number of hypothesis as a function of time and of step index is represented below.



The notations below describe the algorithm :

$h$  is a relative integer such that  $|h| \leq 2^{i-1}$ .

\*  $E(i, 0)$  is not defined

\*  $E(i, h)$  represents the frequency hypothesis :

$$\frac{-1+2|h|}{2^i} \Delta\omega \quad \text{for } h > 0$$

$$- \frac{-1+2|h|}{2^i} \Delta\omega \quad \text{for } h < 0$$

\*  $O(i, h)$  represents the frequency hypothesis  $\frac{2h}{1+2^i} \Delta\omega$ .

Given 2 consecutive complex samples (A and B), one calculates alternatively sets (1,.) and 0(1,.).

Step 1

$$\begin{cases} E(1,1) = A \exp\left(j \frac{\Delta\omega\Delta t}{2}\right) + B \\ E(1,-1) = A \exp\left(-j \frac{\Delta\omega\Delta t}{2}\right) + B \end{cases}$$

$$\begin{cases} O(1,1) = A \exp\left(j \frac{2\Delta\omega\Delta t}{3}\right) + B \\ O(1,0) = A + B \\ O(1,-1) = A \exp\left(-j \frac{2\Delta\omega\Delta t}{3}\right) + B \end{cases}$$

Given 2 consecutive hypothesis sets  $E(i,.)$  and  $O(i,.)$ , one calculates an hypothesis set for step  $i+1$ . Alternatively by (12) and (13). Only positive or null frequency hypotheses are retained, negative frequency expressions are deduced by symmetry.

(12)

$$\begin{cases} E(i+1,1) = E(i,1) \exp j \frac{\Delta\omega\Delta t}{2} + O(i,0) \\ E(i+1,2) = E(i,1) \exp j \frac{3}{2} \Delta\omega\Delta t + O(i,1) \\ E(i+1,3) = E(i,2) \exp j \frac{5}{2} \Delta\omega\Delta t + O(i,1) \\ E(i+1,4) = E(i,2) \exp j \frac{7}{2} \Delta\omega\Delta t + O(i,2) \\ E(i+1,h) = E\left(i, \frac{h+1}{2}\right) \exp j \frac{2h-1}{2} \Delta\omega\Delta t + O\left(i, \frac{h}{2}\right) \\ E(i+1,2^i) = E(i,2^{i-1}) \exp j \left(2^i - \frac{1}{2}\right) \Delta\omega\Delta t + O(i,2^{i-1}) \end{cases}$$

(13)

$$\begin{cases} O(i+1,0) = E(i,1) + O(i,0) \text{ or } E(i,-1) + O(i,0) \text{ "alternate"} \\ O(i+1,1) = E(i,1) \exp j \frac{2^i}{1+2^{i+1}} \Delta\omega\Delta t + O(i,0) \\ O(i+1,2) = E(i,1) \exp j \frac{2^{i+1}}{1+2^{i+1}} \Delta\omega\Delta t + O(i,1) \\ O(i+1,h) = E(i,h+12) \exp j \frac{2^i}{1+2^{i+1}} h \Delta\omega\Delta t + O\left(i, \frac{h}{2}\right) \\ O(i+1,2^i) = E\left(i, 2^{i-1} - \frac{2^i}{1+2^i}\right) \Delta\omega\Delta t + O(i,2^{i-1}) \end{cases}$$

Having obtained with this algorithm a first approximation  $\bar{\omega}$  of  $\omega$ , one subtracts it from the original samples by (14) before applying (10).

(14)

$$\begin{cases} \Delta\theta \leftarrow \bar{\omega}\Delta t \mod 2\pi \\ \theta \leftarrow \Delta\theta \\ \left. \begin{aligned} \phi_i &\leftarrow \phi_i - \theta \mod 2\pi \\ \theta &\leftarrow \theta + \Delta\theta \mod 2\pi \\ t &\leftarrow t + 1 \end{aligned} \right\} > \end{cases}$$

After development of the phase, one has a sequence of numbers  $\phi_i$  perturbed by quasi-additive noise of variance  $\sigma_\phi^2$  given by (9). In order to get simple expressions, it is wise to take a time origin such that  $\sum t = 0$ .

The model of the developed phase variation as a function of time is linear  $\bar{\phi}_i = \bar{\phi}_0 + \bar{\omega}t, \bar{\phi}_0$  and the above SFE algorithm is similar to a discrete Fourier transform and has the same complexity level :  $\log_2 N$  complex operations (rotation plus sum) per input complex sample for an evaluation every time complex samples.

To obtain an evaluation every  $\frac{N}{2}$  complex samples, one has just to repeat the last step, which adds one complex operation per input sample which gives a complexity of  $(1 + \log_2 N)$ .

This method of calculation increases with the rate of intermediate results until it reaches the obvious value  $N$ .

The selected hypothesis is the frequency value  $\bar{\omega}$  corresponding to the biggest modulus.

$\bar{\phi}$  and  $\bar{\omega}$  being unknown are determined through minimisation (least mean square) of :

$$Q = \sum_i (\bar{\phi}_i - \phi_i)^2 = \sum_i (\bar{\phi}_0 + \bar{\omega}t - \phi_i)^2$$

The estimators for  $\bar{\phi}_0$  and  $\bar{\omega}$  are solutions of the 2 equations obtained by cancelling derivatives of  $Q$  with respect to  $\bar{\phi}_0$  and  $\bar{\omega}$ .

$$(15) \quad \begin{cases} \bar{\phi}_0 = \frac{1}{h} \sum_i \phi_i & \text{since } \sum_i t = 0 \\ \bar{\omega} = \frac{1}{h \Delta t} \sum_i t \phi_i & h = \sum_i 1 \end{cases}$$

$h$  is the number of samples used for the estimations. Obviously these estimators are unbiased. To derive the simple expressions of their variances one assumes time decorrelation of angle samples, and introduces the duration  $\tau = h \Delta t$  of estimation.

(16)

$$\begin{cases} \sigma_{\bar{\phi}_0}^2 = \langle (\bar{\phi} - \langle \bar{\phi} \rangle)^2 \rangle = \frac{1}{h \rho} = \left( \frac{\Delta t}{\tau \rho} \right) \\ \sigma_{\bar{\omega}}^2 = \langle (\bar{\omega} - \langle \bar{\omega} \rangle)^2 \rangle = \frac{12}{h(h^2 - 1) \rho (\Delta t)^2} = \frac{12}{\tau^2} \left( \frac{\Delta t}{\tau \rho} \right) \\ \sigma_{\bar{\omega}\bar{\phi}_0}^2 = \langle (\bar{\omega} - \langle \bar{\omega} \rangle)(\bar{\phi} - \langle \bar{\phi} \rangle) \rangle = 0 \quad \text{since } \sum_i t = 0 \end{cases}$$

The last relation of (16) is very important since it proves the decorrelation of estimators of  $\bar{\phi}_0$  and  $\bar{\omega}$ .

## V. SYNCHRONIZATION PERFORMANCE OF SPREAD SPECTRUM SIGNALS

Spread spectrum signaling is the main candidate for future ranging and positioning, but classical methods of despreading and demodulation which realise synchronisation in several steps (chip, bit, message) require a significant delay unless time is known with a high precision. We prove here that one step synchronisation on a known receiver preamble of 30 data bits, even in poor conditions is sufficient to require no continuous transmission.

We consider a known preamble consisting of  $ks$  chips with a duration equal to  $s$  data bits transmission. The mean module of any chip is noted  $a$ , the additive noise has a mean equal to 0 and a variance  $\sigma^2$ , adaptive filtering is implemented and noise samples are decorrelated between distinct chips.

After reconstruction of the carrier with the  $kb$  chips of the known preamble, a distribution of mean  $kba$  and variance  $v = k\sigma^2$  is obtained. In order to decide the preamble presence, it is natural to compare the modulus of this coherent summation to the threshold  $0 < \lambda ks < ks a$ .

Assuming gaussian noise, we can evaluate the conditional probability  $1 - P_d$  of non detection.

$$\begin{aligned} 1 - P_d &= \frac{1}{\sqrt{2\pi k s \sigma^2}} \int_{(1-\lambda)ksa}^{\infty} \exp\left(-\frac{x^2}{2ks\sigma^2}\right) dx \\ &< \frac{1}{\sqrt{2\pi k \sigma^2}} \int_{(1-\lambda)ksa}^{\infty} \exp\left(-\frac{x^2}{2ks\sigma^2}\right) \frac{x dx}{(1-\lambda)ksa} \\ &< \exp\left[\frac{(1-\lambda)^2 k s a^2}{2\sigma^2}\right] \quad \text{when } \frac{2\pi k s a^2 (1-\lambda)^2}{\sigma^2} > 1 \end{aligned}$$

The above inequalities and conditions are better understood by introducing the energy of one bit :

$$E_1 = k \frac{a^2}{2} T$$

and the density of noise power  $N_0 \frac{\sigma^2}{T}$  where  $T$  stands for the bit duration.

$$1 - P_d < \exp\left[-(1-\lambda)^2 \frac{sE_1}{2N_0}\right] \quad \text{when } (1-\lambda)^2 \frac{sE_1}{N_0} > \frac{1}{2\pi}$$

To evaluate the false detection probability  $P_f$ , one needs the statistical law of the modulus of noise of definite variance. To determine this law, one uses the well known differential identity :

$$\left( \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx \right) \left( \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{y^2}{2\sigma^2}\right) dy \right) = \exp\left(-\frac{r^2}{2\sigma^2}\right) \frac{r dr d\theta}{2\pi\sigma^2}$$

$$P_f = \int_{\lambda ks a}^{\infty} \exp\left(-\frac{r^2}{2ks\sigma^2}\right) d\left(\frac{r^2}{2ks\sigma^2}\right) = \exp\left(-\frac{\lambda^2 s E_1}{2N_0}\right)$$

By eliminating the unknown  $\lambda$  between the expression of  $1 - P_d$  and  $P_f$ , one finds the simple inequality (17) which is valid for any data modulation.

$$(17) \quad \sqrt{\ln \frac{1}{1-P_d}} + \sqrt{\ln \frac{1}{P_f}} > \sqrt{\frac{sE_1}{N_0}}$$

To keep  $s$  small, it is advisable to proceed in two steps. In the first step, one takes  $\lambda$  such that  $1 - P_d$  is guaranteed, but with a high  $P_f$ , and the false synchronisations are eliminated in the second step by testing the demodulation likelihood for a few more bits.

## VI. FIXED POINT LOCATION

Well-stabilised geostationary satellites are not completely fixed with respect to the earth surface. For a stabilisation within an angle of 0.1 degree, the diameter of the relative satellite orbit as it appears from the earth is the order of 120 km. To take advantage of this large base for earth positioning, one measures periodically the distance and the angle to the satellite trajectory, and then these intermediary results are combined to estimate the coordinates of the point to be located.

To calculate the intermediate results, one applies equations (15) assuming the satellite trajectory being a straight line during estimation time interval  $\tau$ .

The signal used for these elementary estimations is characterized by a bandwidth occupancy  $B$  (chip frequency of carrier) and a bit rate  $\frac{R}{N}$  ( $\ll B$ ) with an energetical signal to noise ratio  $\rho = \frac{E_1}{N_0} \geq 1$  since :

$$\rho = \frac{A^2}{2\sigma^2} = \frac{RE_1}{2N_0B} = \rho_1 \frac{R}{2B}$$

$$(18) \quad \sigma_{\bar{\omega}}^2 = \frac{2B}{\tau \rho_1 R^2} \quad \sigma_{\bar{\Phi}_0}^2 = \frac{24B}{\tau^3 \rho_1 R^2}$$

These expressions show that estimation performance increases with the spreading factor ( $R/B$ ). If the signal  $s(t)$  is transmitted from the satellite at time  $t$ , at a distance  $D(t)$  function of the same time, one receives  $s\left[t - \frac{D(t)}{C}\right]$ .

During any elementary estimation, it is correct to assume that  $D(t)$  is a linear function of time  $t$  :

$$D(t) = D + tV \cos \alpha$$

In this expression,  $D$  is the distance between the middle of the satellite trajectory used for the elementary estimation and the point  $L$  to be located.  $V$  is the satellite velocity,  $\alpha$  is the unsigned angle ( $0 < \alpha < \pi$ ) between the satellite trajectory and the point direction.

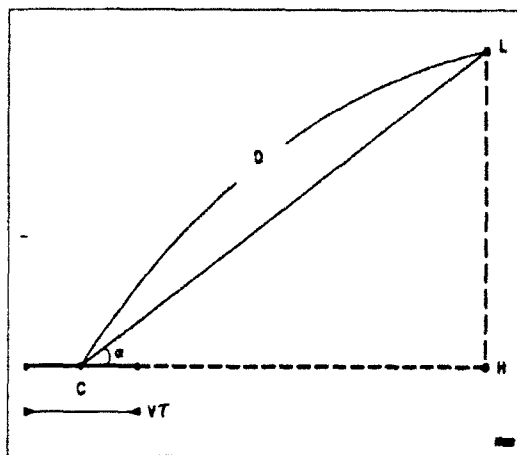


Figure 6 : Geometry of elementary estimations

Carrier or chip frequency in the satellite frame can be represented by the analytic signal  $\exp j(\omega t + \phi)$ . Neglecting relativistic effects, the carrier received in  $L$  is :

$$\exp j\left[\omega\left(t - \frac{D + tV \cos \alpha}{C}\right) + \phi\right]$$

which is identified with :  $\exp j(\bar{\omega}t + \bar{\Phi}_0)$

$$\begin{cases} \bar{\omega} = \omega\left(1 - \frac{V \cos \alpha}{C}\right) & \text{and } \bar{\Phi}_0 = \phi - \frac{\omega D}{C} \bmod 2\pi \\ \text{with } \omega = 2\pi B \end{cases}$$

The satellite trajectory is assumed to be known with perfect accuracy, so that  $D$  and  $\alpha$  are the only parameters which have to be calculated from the estimations of  $\bar{\omega}$  and  $\bar{\Phi}_0$ .

The estimators for  $\bar{\omega}$  and  $\bar{\Phi}_0$  allows us to estimate the length  $D$  of  $CL$  and of its projection  $D \cos \alpha$  on the tangent trajectory in  $C$  after removing phase ambiguity by taking into account an approximation of  $D$ .

The variances of the estimators for  $D \cos \alpha$  and  $D$  are deduced from (18).

(19)

$$\begin{cases} \sigma_{D^2}^2 = \frac{C^2}{2\tau^2 \rho_1 B R^2} & \sigma_{\cos^2 \alpha}^2 = \frac{12C^2}{2\tau^2 \rho_1 B R^2 V^2} \\ \sigma_{D^2}^2 = \frac{\tau^2 V^2}{12} \sigma_{\cos^2 \alpha}^2 & \sigma_{D^2 \cos^2 \alpha}^2 = \left(1 + \frac{12D^2}{\tau^2 V^2}\right) \sigma_{D^2}^2 \end{cases}$$

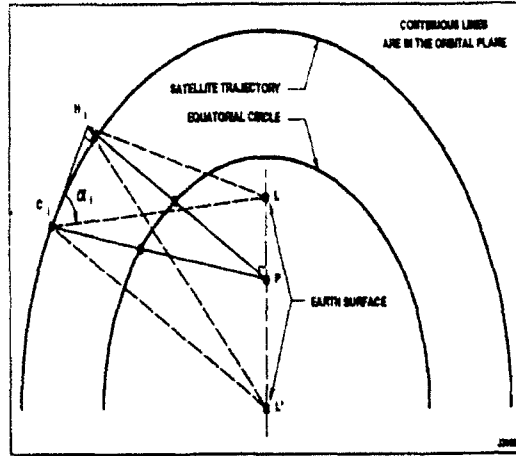


Figure 7 : Notations (low - orbiting satellite)

P is the orthogonal projection of point L on the osculating plane and H is the orthogonal projection of P on the tangent in  $C_i$ .

When the satellite runs at speed V along its trajectory, the sequence of elementary estimations (index i) gives rise to a sequence of lines  $H_iP$ .

Let  $V \cos \gamma_i$  and  $V \sin \gamma_i$  be the coordinates of the trajectory velocity at point  $C_i$ . The equation of the line  $H_iP$  is of the form  $x \cos \gamma_i + y \sin \gamma_i + d_i = 0$  with  $d_i = D_i \cos \alpha_i$  when the axis origin O is the center of the oscillating circle.

The coordinates (x, y) of P minimise the quadratic form :

$$P(x, y) = \sum_i (x \cos \gamma_i + y \sin \gamma_i + d_i)^2$$

Estimator expressions are greatly simplified if the Ox axis is such that :

$$\sum_i \sin \gamma_i = 0$$

which means qualitatively that Ox is parallel to the trajectory part used in the estimation process.

(20)

$$\begin{cases} x = -\frac{\sum_i d_i \cos \gamma_i}{\sum_i \cos^2 \gamma_i} \\ y = -\frac{\sum_i d_i \sin \gamma_i}{\sum_i \sin^2 \gamma_i} \end{cases} \quad \text{with} \quad \sum_i \sin \gamma_i = 0$$

These estimators are unbiased, and their variances are obtained by taking the mathematical expectation of the square of their differential expression, and recalling that :

$$\sigma_{d_i}^2 = \sigma_{D^2 \cos^2 \alpha}$$

(21)

$$\begin{cases} \sigma_x^2 = \frac{\sigma_{D^2 \cos^2 \alpha}}{\sum_i \cos^2 \gamma_i} = \frac{\tau}{T} \sigma_{D^2 \cos^2 \alpha} \\ \sigma_y^2 = \frac{\sigma_{D^2 \cos^2 \alpha}}{\sum_i \sin^2 \gamma_i} = \frac{3}{2} \left( \frac{\tau}{T} \right)^3 (\Delta \phi)^{-2} \tau^2 \sigma_{D^2 \cos^2 \alpha} \end{cases}$$

In the above approximate expressions,  $\tau$  is the elementary measurement duration. T is the global measurement duration along an arc of length  $\frac{T}{\tau} \Delta \phi$  expressed in radian. The above approximations are derived from identities given in [6] p.30.

$$\begin{cases} \sum_{k=1}^n \sin^2 kx = \frac{n}{2} - \frac{\cos(n+1)x \sin x}{2 \sin x} & x = \Delta \phi \\ \sum_{k=1}^n \cos^2 kx = \frac{n}{2} + \frac{\cos(n+1)x \sin x}{2 \sin x} & n = \frac{T}{\tau} \end{cases}$$

For low orbiting satellites, one can assume :

$$\frac{12D^2}{\tau^2 V^2} < 1 \quad \text{so that} \quad \sigma_{D^2 \cos^2 \alpha}^2 \approx \sigma_{D^2}^2$$

(22)

$$\sigma_{D^2 \cos^2 \alpha}^2 \approx \frac{0.05 C^2}{\tau \rho_1 B R^2} \Rightarrow \begin{cases} \sigma_x^2 \approx 0.05 \frac{C^2}{\tau \rho_1 B R^2} \\ \sigma_y^2 \approx 0.15 \frac{C^2}{T \rho_1 B R^2 \left( \frac{T \Delta \phi}{\tau} \right)^2} \end{cases}$$

Typical values are as follows.

$$\begin{array}{lll} \tau = 1s & T = 600s & \Delta \phi = 1.25 \cdot 10^{-3} \\ C = 3 \cdot 10^8 \text{ m/s} & V = 8 \cdot 10^3 \text{ m/s} & \rho_1 = 10 \\ B = 4 \cdot 10^8 \text{ Hz} & R = 100 \text{ bit/s} & \end{array}$$

With these values gives :

$$\sigma_x^2 < 0.3 m^2 \quad \text{and} \quad \sigma_y^2 = 1 m^2$$

This numerical evaluation shows that the geographical precision is fixed only by the carrier frequency and the frequency precision, and the distance to the trajectory earth cross section of the point L to be located.

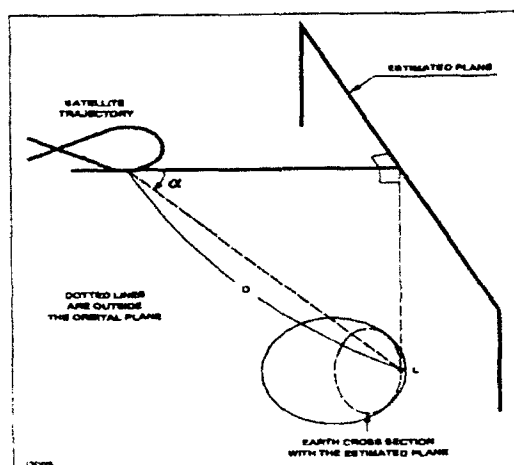


Figure 8 geostationary satellite notations.

In the *geostationary satellite* cases, the trajectory is approximatively parallel to the earth rotation axis and  $\frac{12D^2}{\tau^2 V^2} \gg 1$  which yields the expressions (23).

(23)

$$\sigma_{D^2 \cos^2 \alpha}^2 = 0.6 \frac{C^2 D^2}{\tau^3 \rho_i B R^2 V^2} \Rightarrow \begin{cases} \sigma_{x_1}^2 \approx 0.6 \frac{C^2 D^2}{T \tau^2 \rho_i B R^2 V^2} \\ \sigma_{x_2}^2 \approx 0.9 \frac{C^2 D^2}{T^3 (\Delta \phi)^2 \rho_i B R^2 V^2} \end{cases}$$

As compared with the low orbiting situation, the ratio  $\frac{D^2}{\tau^2 V^2}$  is very unfavourable since  $D = 36 \cdot 10^6$  and  $V$  is of the order of  $3 \text{ m/s}$ .

In order to obtain a useful level of performance, it is necessary to take big values for  $BR^2$  which is possible for example with a of direct digital video broadcasting signal. Inspection of the expressions given by (23) shows that one gets the best results when  $\tau$  is comparable with  $T$ , since  $\Delta \phi$  represents the tangent angle variation during the interval  $T$ .

Typical values are as follows.

$$\begin{array}{lll} \tau = 400 \text{ s} & T = 1200 \text{ s} & \Delta \phi = \frac{1}{32400} \text{ rad} \\ C = 3 \cdot 10^8 \text{ m/s} & V = 3 \text{ m/s} & \\ BR^2 = (40 \cdot 10^6)^3 \text{ Hz}^3 & D = 36 \cdot 10^6 \text{ m} & \rho_i = 10 \end{array}$$

The numerical value given to  $\Delta \phi$  assumes that the direction of arrival of the satellite signal, varies of  $\frac{1}{60}$  radian in 6 hours which is optimistic.

However  $\sigma_{x_2}^2$  is too big to be of any use, but  $\sigma_{x_1}^2 = 600 \text{ m}^2$  is adequate for practical use.

This numerical observation can be explained by the fact that the elementary estimation procedure produces planes perpendicular to the trajectory tangent which gives a well defined intersection only when these planes make between them an angle big enough.

In practice, one can have a good location accuracy by taking into account the estimation of  $D$  which is excellent and the intersection with the geoid. Since the time integration is short compared to the period trajectory, one can evaluate  $\sigma_{D^2}^2$  by replacing  $\tau$  by  $T$  in (19). Except for points situated just under the satellite (*perpendicular to the trajectory plane*), location precision is of the order of  $100 \text{ m}$  after an integration of duration less than 20 minutes.

## VII. PRINCIPLE OF APPLICATION OF PREVIOUS METHODS

In the preceeding section, it has been assumed that the satellite trajectory and especially its plane orientation ( $10^{-7} \text{ rad}$  for geostationary satellite) is known by the receiver L. This information as well as the on-board frequency accuracy can be generated by the satellite.

The trajectory characteristics can be deduced from a known location, by elementary estimation of distance and doppler frequency. Since the estimation time can last several satellite periods (*for a stationary satellite*), one obtains a satellite location precision at least 10 times better than the operational earth location precision which lasts 20 minutes.

Another important application of simultaneous estimation of distances and their derivatives, is the tracking of terrestrial vehicles. In this last case one supposes that the vehicle is also able to estimate its path by independent means (accelerometers for example) for a time interval of the order of 600 s. Provided that the vehicle speed is not too high, the above location method will give the absolute position of the centre of gravity of its estimated local trajectory. The orientation of this local trajectory is done by iterative hypothesis testing. Once the location process is initiated, trajectography can be carried out and the speed limitation condition can be progressively removed.

### VIII. CONCLUSION

It has been demonstrated by statistical and geometrical considerations, that terrestrial location can be obtained with a single geostationary satellite provided that its trajectory is "perfectly" known. The same considerations shows also that satellite tracking (location) can be obtained by signal processing of signals received at a single earth station of known geographical coordinates. However there is a long way to go before an operational system with prescribed characteristics, including time and frequency precision, can be realised.

### DISCUSSION

#### C. GOUTELARD

Votre présentation m'a beaucoup intéressé. Je souhaiterais vous poser deux questions:

1. Pouvez-vous préciser la sensibilité au bruit de cette méthode que vous présentez comme très robuste?
2. Vous auriez pu utiliser une décomposition en ondelettes qui semble bien adaptée aux signaux non stationnaires. Pourquoi ne l'avez-vous pas utilisée?

*I was very interested in your presentation. I would like to pose two questions to you, namely:*

1. *Can you predict the sensitivity to noise of this method which you claim to be very resistant?*
2. *You could have adopted a decomposition in wavelets which appears well adapted to unsteady signals. Why did you not use it?*

#### AUTHOR'S REPLY

1. Il s'agit d'une méthode fondée sur le principe du maximum de vraisemblance. Le niveau de performance souhaité (par exemple sensibilité au bruit) détermine la complexité à mettre en oeuvre pour l'obtenir.
  2. Le modèle de signaux et de ses perturbations sont supposés connus aux valeurs de paramètres près. Ces valeurs sont estimées de façon optimale (moindres carrés par exemple) en utilisant les échantillons entourant les échantillons sur lesquels on veut effectuer les traitements (synchronisation, démodulation, etc.). On s'affranchit ainsi des problèmes d'instabilité dans la mesure où le modèle utilisé est suffisamment général dans la situation considérée.
- 1. This is a method based on the principle of maximum likelihood. The required performance level (e.g. sensitivity to noise) determines the complexity involved in its use.*

2. *The model of the signals and the disturbances are assumed to be known to within the values of the parameters. These values are estimated in an optimal manner (e.g., least squares) using samples around the samples to be processed (synchronization, demodulation, etc.). This eliminates the problems of unsteadiness.*

- [1] "Receiver processor for detection of 406 MHz SARSAT distress message".  
12th AIAA International Communication Satellite Systems Conference.  
P.DUMONT, M.BONNERY,  
A.MARGUINAUD.
- [2] "Demodulation of messages received with low signal to noise ratio".  
ICDCS 24-28 April 1989.  
A.MARGUINAUD, T.QUIGNON,  
B.ROMANN.
- [3] VILLE (J) "Signaux Analytiques à spectre borné" - Câbles et transmission 4:9 (1950).
- [4] CHENEY (E.W) "Introduction to approximation theory - Chelsea Publishing Company (N.Y)".
- [5] ALCATEL TECHNICAL NOTE n°313.  
A.MARGUINAUD
- [6] GRADSHETEYN (IS) and RYZHIK (IM)  
"Tables of integrals, series and products  
Academic Press".

# RADIOLOCATION IN THE LOWER ELF FREQUENCY BAND

C.P. Burke

D. Llanwyn Jones

Physics Department, King's College,  
London, U.K. WC2R 2LS

## SUMMARY

A system for recording Extremely Low Frequency (ELF) noise bursts produced by global lightning activity is described.

The location of the source lightning flashes for some 320 such events has been deduced from these data. The range of the sources was found by modelling the data using a least-squares fit to the complex wave impedance obtained from the standard propagation theory applicable to the Earth-ionosphere spherical-shell waveguide. The source bearings were deduced from two orthogonal components of the magnetic field vector using the usual goniometric technique. The source range and bearing serve to locate each source on the surface of the Earth.

The data show that the majority of ELF event sources are located in tropical regions, an average of 7.4 Mm away from the observing station situated at 51.14 degrees North, 1.44 degrees West (Geographic). The use of such a system to monitor worldwide thunderstorm activity, so far as this relates to ELF events has been demonstrated.

## 1. INTRODUCTION

The cavity formed by the surface of the Earth and the lower fringe of the ionosphere acts as a spherical-shell electromagnetic waveguide, and at the lower end of the ELF band (5-45 Hz), the attenuation of waves within the cavity is low (typically 0.1 dB/1000 km at 10 Hz - Jones [1]). An effect of the low propagational attenuation is that the 'Q-factor' of the natural electromagnetic cavity resonator formed by the Earth-ionosphere system is of order 5. The resonance modes of this cavity are excited by global lightning activity - these are the so-called Schumann resonances.

In this paper we are concerned with the location of the source of large amplitude, Extremely Low Frequency (ELF) radio pulses whose origin is lightning flashes. These pulses are termed ELF events. The amplitude of usable ELF events is, typically 5 - 10 times the rms amplitude (1mV/m) of the Schumann resonances when observed in the band 5 - 45 Hz. The occurrence rate of distinguishable events is one to two per minute.

Jones and Kemp [2] demonstrated that the range and bearing of the source lightning flash of an ELF event could be determined from the Fourier frequency spectrum of either the electric or the magnetic components of the received signal.

The electric and magnetic field spectra however depend upon the (unknown) spectrum of the source lightning flash. An improved technique is that of Kemp and Jones [3] who based their analysis upon the wave impedance of the signals. This method was further developed by Ishaq and Jones [4]. The wave impedance is simply the (complex) ratio of the electric and magnetic field frequency spectra.

The use of a wave impedance method for deducing the range of a source was first suggested by Wait [5] who considered wave propagation in a planar Earth-ionosphere waveguide. Our method, however, employs the theory for a spherical-shell waveguide in which ELF signals make multiple circulations of the globe. This characteristic is the essential feature of the present method of range determination which is only weakly dependent on the ionospheric model. The seminal theoretical work for circum-global ELF propagation is that of Wait [6] and Galejs [7] who follow the early pioneering work of Sommerfeld [8].

The geomagnetic field produces anisotropy in the ionospheric plasma which, in principle, leads to non-reciprocal propagation. This feature was investigated by Pappert and Moler [9] employing a cylindrical geometry. They concluded that non-reciprocity was unlikely to be observable at ELF. The same conclusion was also reached by Behroozi-Toozi and Booker [10] in a comprehensive theoretical study of ELF propagation.

## 2. INSTRUMENTATION

At ELF only two field components are radiated effectively by a vertical electric dipole which is used here to model a lightning flash, these fields being a radial electric field  $E_r$  and a horizontal magnetic field  $H_\phi$ . Horizontal dipoles radiate ineffectively at ELF - Wait [6]. The accuracy of representing a lightning flash by a Hertzian dipole at ELF and large source-observer distances is very good - LeVine and Meneghini [11].

To effect the analysis of ELF events, signals were detected by receiving antennas and passed through amplifying and filtering electronics. After digitization using a transient waveform recorder, signals were stored on floppy disc. A schematic of the ELF receiving and recording equipment is given in Fig.1.

The electric field antenna consists of a tripod 2m high. This antenna, which has been used in previous investigations e.g. Jones and Kemp [2], is essentially a charge sensing device whose output voltage  $V_e$  due to an applied field  $E$  is given by  $V_e = E \times h_{eff}$ , where  $h_{eff}$  is the effective



tive height of the antenna, the determination of which is described later. A tripod construction was chosen since this gives good mechanical stability. Stability is important since vibrations of the antenna in the Earth's natural electrostatic field caused, for example, by the wind can result in large unwanted signals. A separate parallel plate antenna was used for calibration purposes. This consists of two square metal plates of side 1m by 1m, separated by 0.1m.

Because the magnetic field vector lies in the horizontal plane, the output signals from two orthogonal solenoidal antennas define the received field. Each antenna consisted of 54,000 turns of copper wire distributed over six bobbins and threaded on to a ferro magnetic core of measured effective permeability 530. A secondary calibration coil of ten turns was wound over the main bobbins, the use of which is described later. A grounded aluminium screen encased this arrangement (to provide screening against electric fields) and the final assembly was enclosed in plastic pipes for protection. To reduce interference from 50 Hz power lines, the antennae were located in a rural location at 51.14 degrees North, 1.44 degrees West. The magnetic antennae were accurately aligned north-south and east-west (geographic) by using local landmarks.

Fig. 2 shows the electric amplifier design. A high impedance input stage was built around an 071 operational amplifier which is a high impedance chip whose noise was negligible compared to the external sources. An input resistor of 10G Ohms was used as this gave a relatively flat response for this stage down to 0.3 Hz. The response of the stage drops off at low frequencies because the antenna acts as a voltage source in series with a capacitor (of the order of 75 pF) whilst the 10 G Ohm resistor has a stray capacitance across it of the order of about 10 pF. The circuit thus behaves as a high pass filter with a cut-off frequency of 0.3Hz. The 150 k Ohm resistor in series with the antenna serves to attenuate in-band intermodulation products from high frequency transmissions such as television signals.

This initial preamplifier was contained in its own sealed and screened box that connected directly to the tripod antenna. It was connected to the main preamplifier by two leads: one for power and one for signal. Some initial preamplification then occurs before the signal passes through a 4-pole, 0.1dB ripple, Chebyshev high pass filter with a band edge of 5 Hz (the 'wind filter') and then passes through the power-line-frequency rejection filter, the EF50, a commercial active filter. A 3-pole, 0.1dB ripple, low pass Chebyshev filter follows this with a band edge of 55.5 Hz. This was used to attenuate power-line harmonics signals. The signal is then split to give a balanced output over the relatively long distance to the recording hut.

The magnetic system is shown in Fig. 3. Here, a high input impedance was not the dominating concern. Instead, because the input signal voltage was expected to be very small, it was necessary to have a very low noise amplifier. Two FETs in parallel were used as the input stage and following this there is a buffer built around an OP-27 to prevent loading of the FETs by the following electronic integrator. A low pass passive filter is incorporated as in the electric system. The integrator gives a good 1/f response down to below 5 Hz. After this, the signal passes through similar stages to the electric system (EF50, power frequency harmonic filter and balanced output stage).

Because the signals were to be analysed in the frequency domain, antialiasing filters were required. To utilize the full dynamic range of the Analogue-to-Digital convertor (ADC), the aliased signal had to be 72 dB below the base band signal at the Nyquist frequency (156 Hz). The filters consisted of a 60 dB attenuation commercially-manufactured part with a further 12 dB of attenuation being supplied using 3-pole elliptic filters. Four such filters were employed: one each for the electric field (E), the North-South and East-West magnetic field components (Hns and Hew) and a fourth for the trigger channel (T). The trigger channel was derived from the E channel (as this is omni directional) and incorporated an additional 3-pole high pass Butterworth filter: of band edge 13 Hz to prevent spurious triggering of the ADC caused by the wind and atmospheric conductivity flicker noise, which was apparent despite the earlier wind filter.

The ADC (Data Laboratories model DL1200) was a 12-bit, microprocessor controlled, 4-channel instrument designed to capture transient signals. A major feature of the instrument was the full remote control of all features via the GPIB interface.

The DL1200 enters its recording mode immediately it is 'armed'. The input signals are sampled and the oldest data is discarded as the new data replaces it (somewhat like a continuous loop of tape). This means that pre-trigger information is retained. When the trigger occurs, there is already information in the memory. The DL1200 enabled all four channels to be sampled synchronously and digitised each waveform to give 8192 12-bit words sampled at 200  $\mu$ S intervals. A microcomputer was used to remotely control the DL1200 via the GPIB.

Waveforms captured by the DL1200 on receipt of a trigger signal were viewed on an oscilloscope screen; if considered to be of sufficient quality, the digitised waveforms were transferred to the computer by the GPIB and dumped as a binary file to floppy disc. In the disc file only 1 in 16 (i.e. 512) of the original data samples were retained thus increasing the effective sampling interval to 3.2ms and reducing the Nyquist frequency to 156.25 Hz.

Both the electric and magnetic systems required calibration. Two calibrations were used for the magnetic systems. The secondary (local) calibration used the small coils wound on to the main aerial bobbins. By injecting pulses into these coils and comparing the output of the receiving system with the input signal in the frequency domain using a spectrum analyser, the frequency response of both magnetic systems could be determined. The absolute sensitivities of each antenna were measured by creating a known magnetic field around each aerial and noting the output response. This known field was produced by the amplified signal from a sine wave oscillator driving a remote calibration coil of some 100 turns of wire which was located some tens of metres from the receiving solenoid. A sine wave was chosen rather than the broadband signals used for the secondary calibration since the broadband signals gave a poor signal to noise ratio in this case. The mid-band absolute sensitivity of the magnetic field systems (measured at the input of the ADC) was about 3 V/nT.

For the electric-field system it was not possible to create a uniform electric field about the antenna due to its size so a different approach was adopted. Firstly, the frequency response of the electronics was established by using a spec-

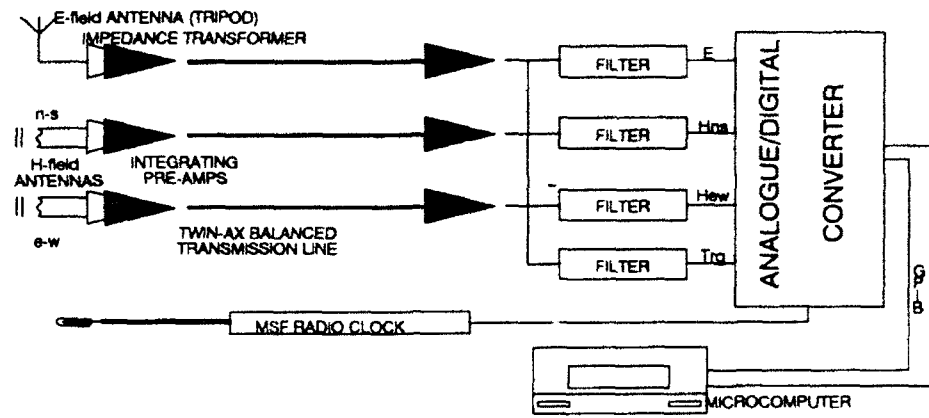


Fig. 1 The ELF receiving system

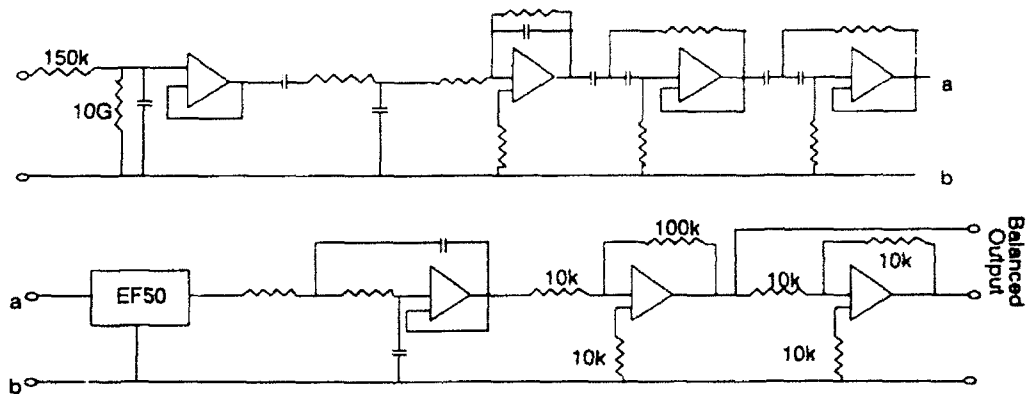


Fig. 2 The Electric Preamplifiers

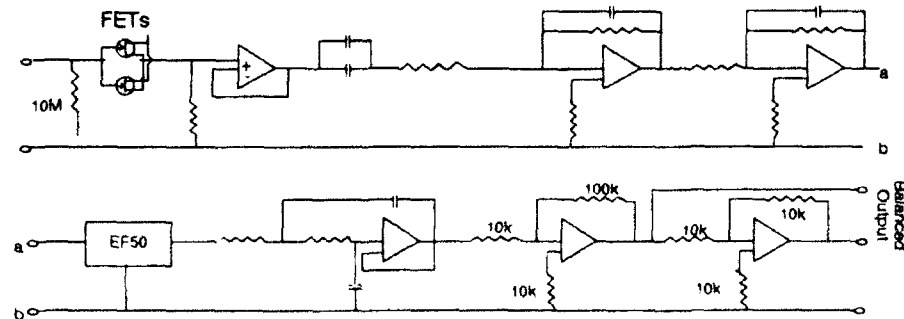


Fig. 3 The Magnetic Preamplifiers

trum analyser. The system was driven through a 75 pF capacitor dummy antenna to represent the known aerial capacitance. The effective height of the antenna was determined by comparing the signal received by the tripod antenna with that of the parallel plate antenna from the 16 kHz Rugby V.L.F. (GDR) transmitter. Here the effective height of the parallel plate antenna is well known. In this way, the effective height of the tripod antenna was estimated as  $1.01 \pm 0.03\text{m}$  and the overall absolute sensitivity of the system was about 7m (i.e. 7 V/V/m).

Figs. 4-6 show the amplitude and phase characteristics of the 3 data channels ( $E$ ,  $H_{ns}$  and  $H_{ew}$ ). The amplitude response is reasonably flat in the band of interest, showing, in the case of the magnetic systems, that the integrating amplifiers perform well. The effects of the EF50 notch, wind and power-line frequency harmonic filters can also be seen.

Fig. 4 Amplitude Frequency Response of the Electric System (Incl. Filters)

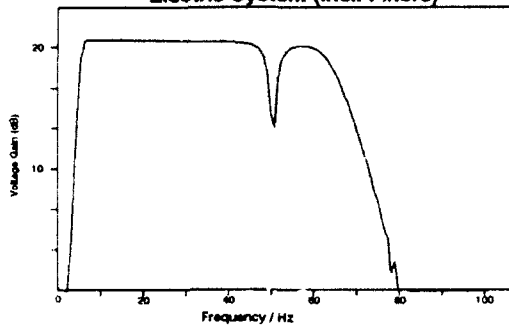


Fig. 5 Magnetic System 1 Amplitude Transfer function

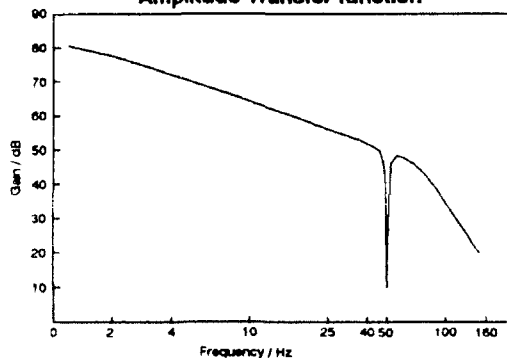
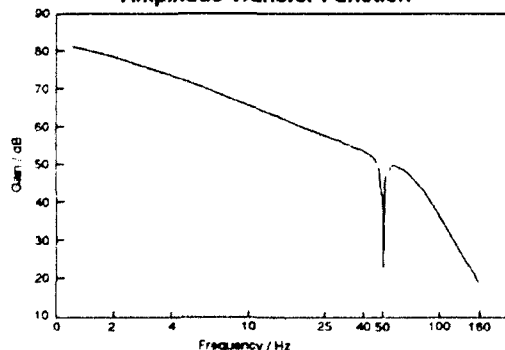


Fig. 6 Magnetic System 2 Amplitude Transfer Function



### 3. DATA ANALYSIS THEORY

To locate the source lightning flash, we require to know the source bearing and the source range (propagation distance). In this section we summarize the necessary propagation theory needed to estimate the source range,  $d$ .

The radial electric field  $E_r$  and azimuthal magnetic field  $H_\phi$  due to a vertical electric dipole of current moment  $I ds$  at a distance  $d = a\theta$  from the observer is given to good accuracy by:

$$E = i \frac{I ds(\omega) \nu(\nu+1) P_\nu^0(-\cos \theta)}{4\pi a^2 \epsilon_0 \omega h \sin(\pi \nu)} \quad (1)$$

$$H = \frac{-I ds(\omega) P_\nu^1(-\cos \theta)}{4a h \sin(\pi \nu)} \quad (2)$$

(Wait [6], Jones and Kemp [2]) where  $\theta$  is the geocentric angle between the source and the observer,  $h$  is the effective ionospheric height,  $P_\nu^0(-\cos \theta)$  and  $P_\nu^1(-\cos \theta)$  are Legendre functions of (complex) degree  $\nu$  and order 0,1,  $\omega$  is the angular frequency,  $\epsilon_0$  the permittivity of free space and  $a$  is the Earth radius. The wave impedance

$$Z = -i \frac{\nu(\nu+1) P_\nu^0(-\cos \theta)}{a \epsilon_0 \omega P_\nu^1(-\cos \theta)} \quad (3)$$

is thus independent of the source spectrum  $I ds$  (Kemp and Jones [3]). The so-called modal eigenvalue  $\nu$  is related to the attenuation and phase velocity of the radiation in the Earth-ionosphere cavity. We use our experimental data to determine  $\nu(\omega)$  as described below.

Kemp and Jones noted that, when plotted as functions of frequency, both the modulus and argument of the wave impedance,  $\text{Mod } Z$  and  $\text{Arg } Z$  are of the form of a decaying quasi-sinusoidal oscillation the period of which is characteristic of the propagation distance. They went on to derive an asymptotic relation between the propagation distance and the phase velocity,  $v$ . Building upon this, Bliokh et al [12] suggested that Fourier analysis could be used to determine the characteristic period. They considered the Fourier cosine transform of  $|Z|$  i.e.

$$\text{Re} \frac{1}{\pi} \int_{-\infty}^{\infty} |Z(\omega)| e^{-i\omega t} d\omega \quad (4)$$

The maximum of the transform, when plotted as a function of inverse frequency  $\tau$  is a measure of the characteristic period of  $|Z|$ . Numerical experiments in the course of this investigation have shown that the Fourier transform rather than the Fourier cosine transform gives more reliable results. The simple extension to include  $\text{Arg } Z$  has also been made. The periods of  $|Z|$  and  $\text{Arg } Z$  are thus found from the maxima of

$$\text{Mod} \int_{-\infty}^{\infty} |Z(\omega)| e^{-i\omega t} d\omega \quad (5)$$

and

$$\text{Mod} \int_{-\infty}^{\infty} \text{Arg } Z(\omega) e^{-i\omega t} d\omega \quad (6)$$

Using the propagation values deduced by Jones [1] from

Schumann resonance measurements and by applying a least squares fit to a plot of  $\tau_{max}$  versus propagation distance gives

$$d = -125.5\tau_{max} + 20.5 \quad (8)$$

for  $|Z|$  and  $\text{Arg } Z$  where  $d$  is measured in Mm (1000 km) and  $\tau_{max}$  in seconds. This value of the propagation distance is slightly dependent upon the propagation model used in the numerical experiments described above. To refine this estimate, we fit a curve defined by the wave impedance equation (equation 3) to the experimentally measured wave impedance  $Z_e$ . We thus attempt to minimize

$$R = \sum_i |(Z_e(\omega_i) - Z(\nu, \omega_i))|^2 \quad (9)$$

where  $\nu(\omega)$  is modelled by

$$\nu = c_0 + c_1\omega^{c_2} + c_3\omega^{c_4} + i(c_5 + c_6\omega^{c_7} + c_8\omega^{c_9}) \quad (10)$$

and

$$\theta = c_{10} \quad (11)$$

using the simplex algorithm of Nelder and Mead [13]. The propagation data of Jones [1] provides the first estimates of the fitting parameters  $c_0$  to  $c_9$  while the value of  $d$  from the Fourier analysis of  $|Z|$  and  $\text{Arg } Z$  provides the first estimate of  $c_{10}$ . To perform the calculation of  $Z$  (equation 3) we require methods to calculate Legendre functions of complex degree. The rapid and accurate computation of these has been discussed elsewhere (Jones and Joyce [14], Jones and Burke [15]).

After the curve fitting process has terminated the residuals

$$r_i = |Z_e(\omega_i)| - |Z(\omega_i)| \quad (12)$$

are plotted as a function of frequency. If these exhibit no trend, then the curve fit is considered successful, if not, further coefficients are included in the expression for  $\nu$ .

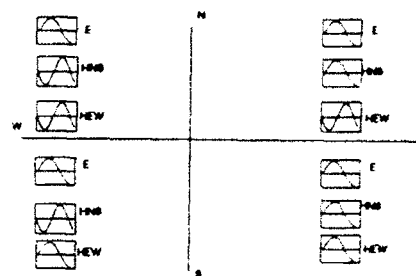
#### 4. RESULTS

Fig. 7 shows the responses of the magnetic-field solenoids to test signals from the remote calibration antenna when this is moved through the four geographic quadrants surrounding the receiving antennas. For a given set of ELF event waveforms,  $H_{ns}$  and  $H_{ew}$  were examined to decide in which quadrant the source was located. It is also necessary to note the polarity of the electric field waveform to resolve the inherent ambiguity of bearing deduced from the magnetic field waveforms. Then, in the frequency domain, the source bearing

$$\beta = \tan^{-1} \left( \frac{H_{ns}(\omega)}{H_{ew}(\omega)} \right) \quad (13)$$

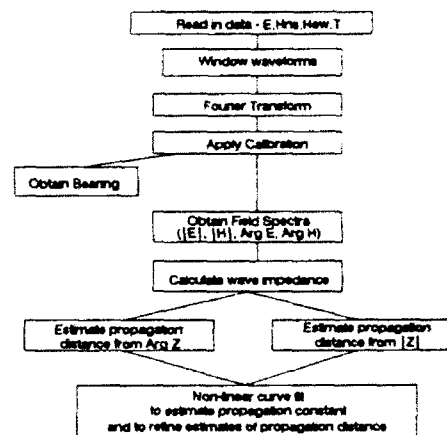
was calculated and a check made to ensure that the source lay in the expected quadrant.

Fig. 7 Aerial Responses



Many hundred events were recorded in the period September 1988 to August 1989 of which some 323 were fully analyzed. Data analysis was conducted in an interactive session using the microcomputer's graphical capabilities. The data analysis is summarized in Fig. 8. Results of the analysis dealing with the modal eigenvalue  $\nu$  (attenuation and phase velocity) are dealt with by Burke and Jones [16].

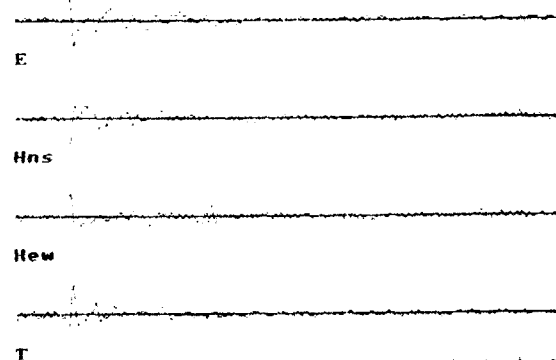
Fig. 8 Data analysis procedure



Consider the waveforms illustrated in Fig. 9 which shows an ELF event captured on all four ADC channels (E, Hns, Hew and T). These waveforms were recorded at 14:04pm on 21st September 1988. The event was preceded and followed by a period of normal, quiet Schumann resonance noise which suggests that the event was caused by an unusually large lightning flash located at a specific point.

Fig. 9 An ELF event

CH FILE E89L14 04.22E CPU EPOCH = 880921



The markers at the foot of the diagram show 100 mS intervals. All the field components show a rapid excursion to mark the beginning of the event; this is followed by a decaying 'ring'. The initial impulse is the signal that has travelled to the recording station via the direct, or shortest route. This is followed by a pulse having a more prolonged excursion of the opposite polarity arriving some 150 mS later. This is the pulse which has travelled on the long great circle path to the recording station.

From the polarity tests described above, it is clear that the polarity of the electric field and the two magnetic field polarities at the beginning of the event locate the lightning flash in the south-west quadrant. Furthermore, by computing the bearing  $\beta$  from equation (13) (which gives the angle of the wavefront from the north-south direction), this bearing is deduced to be  $220 \pm 7$  degrees.

In Figs. 10a and 10b the variations of  $|E_r|$  and  $|H_\phi|$  with frequency are given. The spectra of the two field components are clearly different. However, the Schumann resonance modes can be observed. In the case of the electric field, the odd modes are weak while the even modes are strong. For the magnetic field, the opposite is true; the even modes are weak and the odd modes are quite pronounced. This behaviour is characteristic of spectra from a distance of about 10 Mm. In Figs. 11a and 11b the spectra for  $|Z|$  and  $\text{Arg } Z$  for this event are shown together with the fitted curves deduced from the numerical optimization algorithm. The spectra have been normalized to the wave impedance of free space. The spectra are of the form of a quasi-sinusoidal decaying oscillation. The mean level of  $|Z|$  is a somewhat greater than one and that of  $\text{Arg } Z$  is about 180 degrees. The propagation distance deduced from the optimisation curve fitting process is 9.8 Mm.

Fig. 10a E-Field Amplitude Spectrum

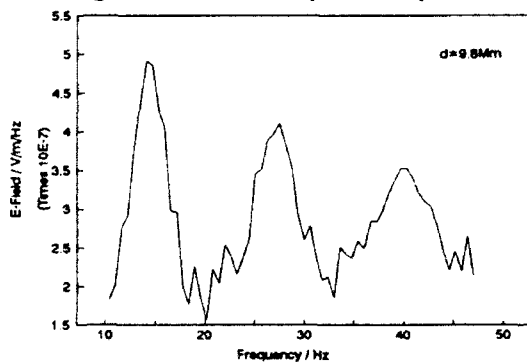


Fig. 10b H-Field Amplitude Spectrum

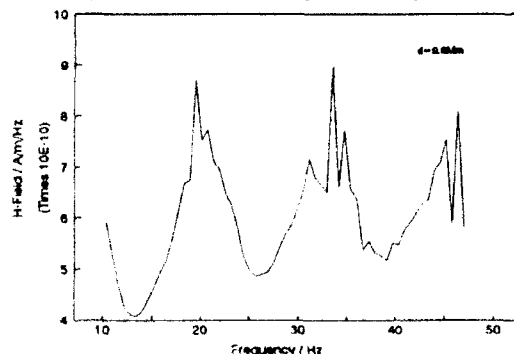


Fig.11a Experimental and Theoretical Wave Impedances (Modulus)

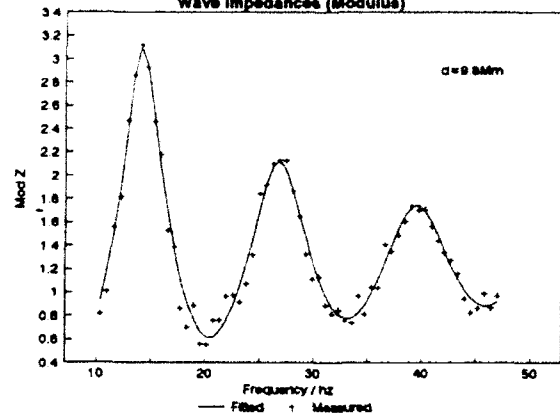
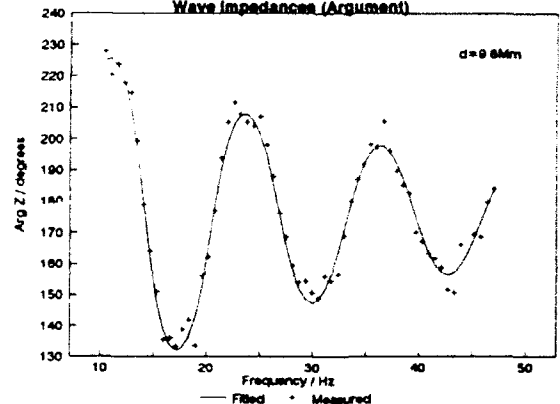
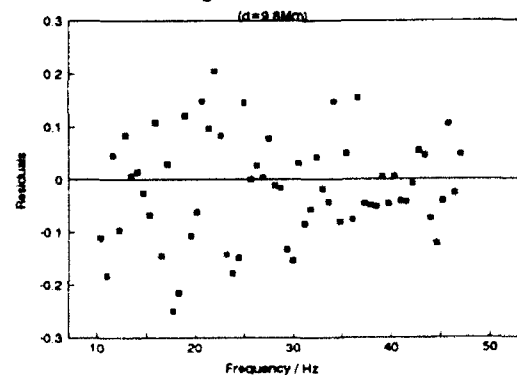


Fig. 11b Experimental and Theoretical Wave Impedances (Argument)



The residuals  $r_i = |Z_{e,i}| - |Z_i|$  from the curve fitting process are shown in Fig. 12. No trend is apparent so it is evident that the optimisation is adequate.

Fig. 12 The Residuals



Figs. 13a and 13b and 14a and 14b show two further examples of the wave impedance spectra of two events with deduced propagation distances of 10.8 Mm and 11.7 Mm respectively. The form of the spectra is similar to the event described above although the 'periods' of  $|Z|$  and  $\text{Arg } Z$  are correspondingly greater.

Fig. 15 shows a histogram of source distances for all the data analysed. The distribution, which is clearly not Gaussian, has a mean of 7.4 Mm and a standard deviation of 2.3 Mm.

Of somewhat more interest is Fig. 16 which shows the geographic locations of all the ELF events analysed. Each point represents one ELF event which has been located using the technique described above. It is clear that the sources of ELF events tend to be located in tropical regions. This presumably reflects the well-known fact that world thunderstorm activity is a maximum in the tropics. Few events occurred north of the observation site (in northern polar regions or northern Pacific regions) which is due to the paucity of lightning activity in these areas.

Fig. 13a Experimental and Theoretical Wave Impedances (Modulus)

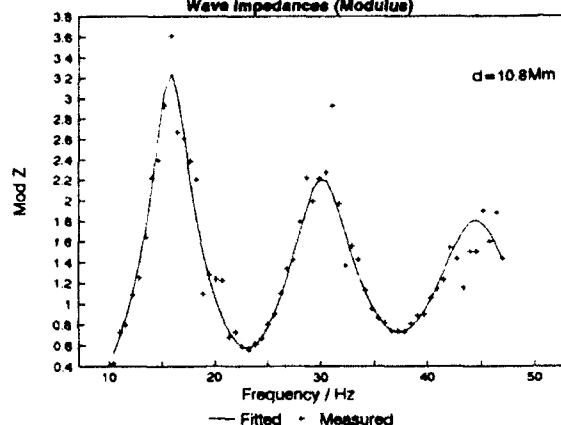


Fig. 14a Experimental and Theoretical Wave Impedances (Modulus)

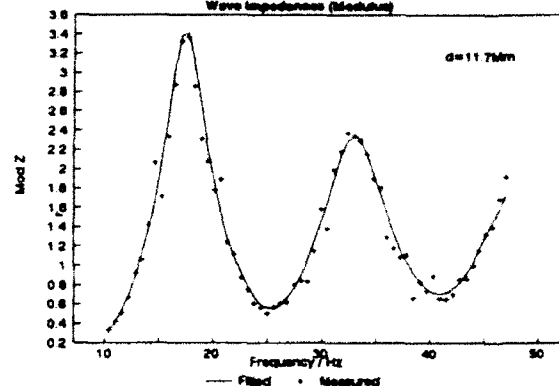


Fig. 14b Experimental and Theoretical Wave Impedances (Argument)

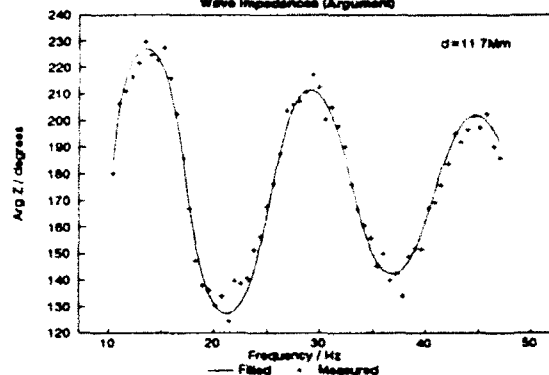


Fig. 13b Experimental and Theoretical Wave Impedances (Argument)

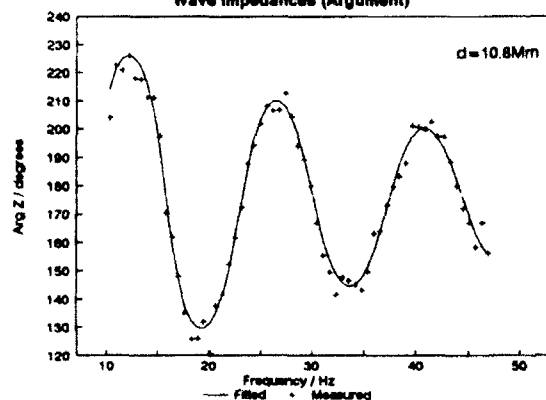


Fig. 15 Histogram of Propagation Distances

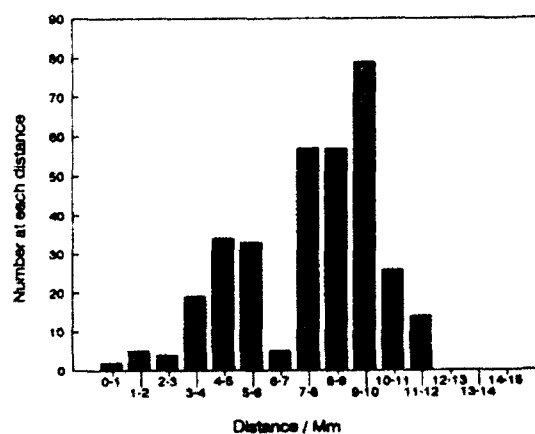
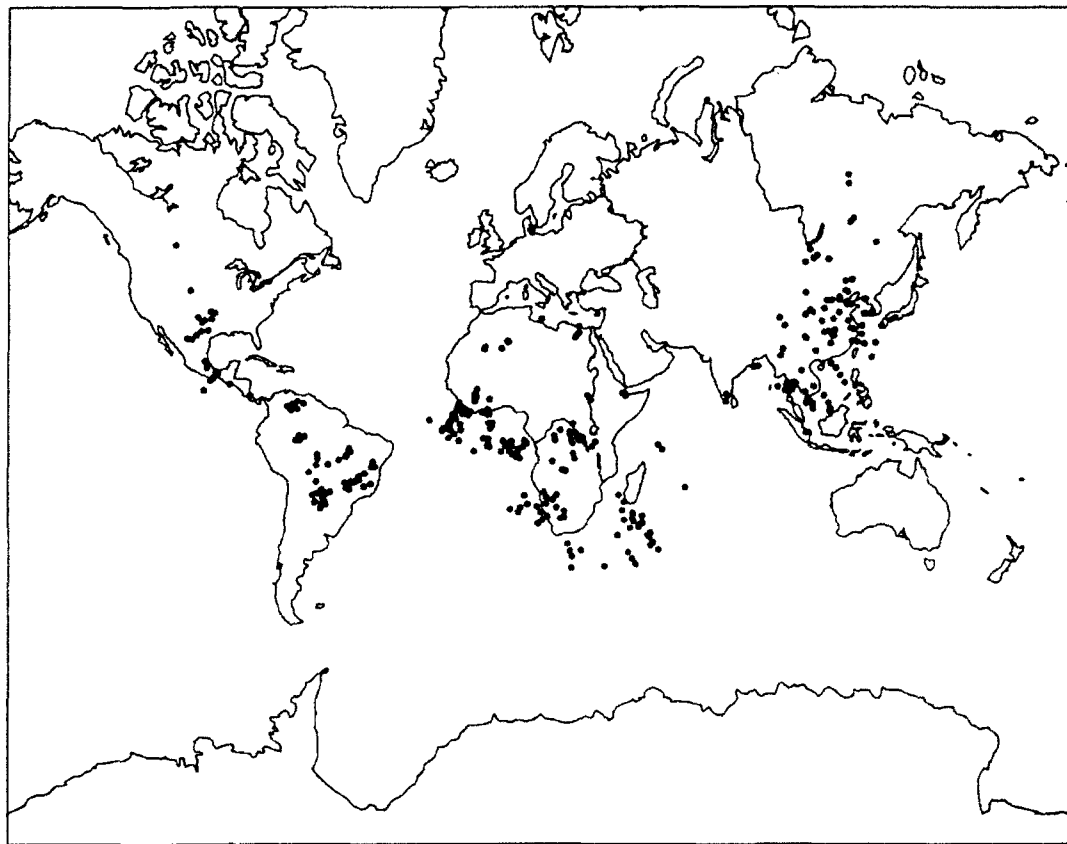


Fig. 16



## 5. CONCLUSION

A system for the reception and recording of ELF events has been described. Some 320 events have been analysed and located using a curve fitting optimisation algorithm. Most were located in tropical regions. A method of monitoring worldwide thunderstorm activity from a single station (so far as it relates to ELF events) has thus been demonstrated.

## Acknowledgements

The work described here has been supported by the U.K. Science and Engineering Research Council under grant SGD 10351. We wish to thank Dr. J. Norbury and Dr. R. Morgan for allowing us use of the Rutherford Appleton Laboratory site at Chilton. M. Cramoisan was involved in the design and construction of some of the electronic systems.

## References

- [1] Jones D.LI. Schumann resonances and ELF propagation for inhomogeneous isotropic ionosphere profiles, *J. Atmos. Terr. Phys.*, 29, pp1037-1044, 1967
- [2] Jones D.LI. and Kemp D.T. Experimental and theoretical observations of Schumann resonances, *J. Atmos. Terr. Phys.*, 32, pp1095-1108, 1970
- [3] Kemp D.T. and Jones D.LI. A new technique for the analysis of transient ELF electromagnetic disturbances within the Earth-ionosphere cavity, *J. Atmos. Terr. Phys.*, 33, pp567-572, 1971
- [4] Ishaq M. and Jones D.LI. Method of obtaining radiowave propagation parameters for the Earth-ionosphere duct at ELF, *Electron. Lett.*, 13, pp254-255, 1977
- [5] Wait J.R., Influence of Source Distance on the Impedance Characteristics of ELF Radio Waves, *Proc. I.R.E.*, 48,(7),1960,pp1338-1339.
- [6] Wait J.R. 'Electromagnetic Waves in Stratified Media', (2nd Ed.), Pergamon Press, New York, 1970
- [7] Galejs J. 'Terrestrial Propagation of Long Electromagnetic Waves', Pergamon Press, New York, 1972
- [8] Sommerfeld A.N. 'Partial Differential Equations in Physics', Academic Press, New York, 1949
- [9] Pappert R.A. and Moler W.F. Propagation Theory and Calculations at Lower Extremely Low Frequencies (ELF) *I.E.E.E. Trans. Comm. Electron.*, COM-22, pp438-451, 1974
- [10] Behroozi-Toosi A.B. and Booker H.G. Application of a

simplified theory of ELF propagation to a simplified worldwide model of the ionosphere. *J. Atmos. Terr. Phys.* 42, pp943-974, 1980

[11] Levine D.M. and Meneghini R. Electromagnetic fields radiated from a lightning return stroke: application of an exact solution to Maxwell's equations, *J. Geophys. Res.*, 83, pp2377-2384, 1978

[12] Bliokh P.V., Nickolaenko A.P. and Filippov Yu. F. 'Schumann Resonances in the Earth-ionosphere Cavity', Peter Peregrinus, London, 1980

[13] Nelder J.A. and Mead R. A simplex method for func-

tion minimization, *Comp. J.*, 7, pp308-313, 1965

[14] Jones D.Ll. and Joyce G.S. The computation of ELF radio wave fields in the Earth-ionosphere duct, *J. Atmos. Terr. Phys.*, 51, pp233-239, 1989

[15] Jones D. Ll. and Burke C.P. Zonal harmonic series expansion of Legendre functions and associated Legendre functions, *J. Phys. A: Math. Gen.*, 23, pp3159-3168, 1990

[16] Burke C.P. and Jones D.Ll. An experimental investigation of ELF attenuation rates in the Earth-ionosphere duct, *J. Atmos. Terr. Phys.*, 54, pp243-250, 1992

## DISCUSSION

### J. AARONS

Have you determined the occurrence or the morphology of the events in the equatorial region as a function of season? (You directed your answer towards the propagation rather than the generation).

### AUTHOR'S REPLY

The ELF event activity recorded, which relates to unusually large flashes, agrees well with the known temporal and spatial variation of thunderstorm centers.



# LES SEQUENCES W. G. SEQUENCES BINAIRES QUASI PARFAITES

## THE W.G. SEQUENCES QUASI PERFECT BINARY SEQUENCES

C. GOUTELARD

Laboratoire d'Etude des Transmissions Ionosphériques

Université Paris-Sud

Bâtiment 214, 91405 Orsay Cedex, France

### RESUME -

La recherche de séquences présentant de bonnes fonctions de corrélation périodiques est un sujet sur lequel se focalisent, depuis longtemps, bien des recherches. Les séquences binaires offrent un intérêt particulier à cause de leur simplicité de génération et de traitement.

Les séquences W.G. (Wolfmann-Goutelard) présentées dans cet exposé possèdent des fonctions de corrélation périodiques quasi parfaites, telles qu'en dehors du pic central, ces fonctions sont nulles excepté en un point.

Il est montré alors que la structure des séquences W.G. est unique et qu'elle résulte de l'entrelacement de trois séquences périodiques.

Les propriétés de ces séquences sont étudiées et il est montré que des algorithmes de détermination réduisent la complexité de leur recherche.

Les conditions d'existence - ou de non existence - de ces séquences sont données et les séquences existant pour les longueurs  $N = 8\alpha + 4$ ,  $N = 16\alpha + 4$  et  $N = 32\alpha + 4$  où  $\alpha \in \mathbb{Z}$  sont données jusqu'à  $N = 20000$ .

### I. - INTRODUCTION -

L'utilisation de séquences périodiques ou apériodiques à fonction de corrélation nulle en dehors du pic central est d'un grand intérêt dans les problèmes de radiolocalisation de télécommunication et de radar.

On demande, la plupart du temps, que le signal soit à amplitude constante et que les séquences soient binaires pour faciliter leur génération et le traitement du signal à la réception. Des séquences apériodiques, telles les séquences de Golay, présentent ces caractéristiques (figure 1). Elles nécessitent une émission discontinue gênante dans bien des applications. Des séquences périodiques polyphases [1] [2]

ou à trois moments [3] présentent de telles caractéristiques. On ne connaît pas de séquences binaires présentant ces caractéristiques et il a été montré qu'il n'en existait pas pour les longueurs  $N$  comprises entre 4 et 12100 [4]. Seules les séquences pseudo-aléatoires ou les séquences pseudo-orthogonales [5] ont des fonctions de corrélation s'approchant de celles souhaitées.

Il a été recherché des séquences binaires telles que leur fonction de corrélation demeure nulle sur un espace aussi grand que possible centré sur le décalage nul. Ces séquences nommées Cazac ont été décrites par R.J. Polge [6] qui donne une amorce de procédure de construction limitée, par sa complexité, aux séquences de faibles longueurs. Une recherche exhaustive par ordinateur faite par R. Alexis [7] jusqu'à  $N = 30$  a montré que des séquences présentaient des fonctions de corrélation nulles excepté pour un décalage nul et pour un décalage égal à  $N/2$ . De plus, l'auteur a montré que  $N$  ne pouvait être qu'un multiple de 4.

J. Wolfmann [8] s'est attaché à rechercher les caractéristiques générales des séquences dont les fonctions de corrélation sont nulles excepté pour un décalage nul et un décalage égal à  $N/2$ . Il en a défini les caractéristiques générales et a observé sur des résultats expérimentaux établis sur des séquences de petites longueurs la répétition d'une structure caractéristique du vecteur, défini par le produit de Cauchy de la séquence et de son translaté d'un symbole.

En conservant cette structure pour des longueurs plus grandes il a pu déterminer, par une recherche informatique, des "séquences à auto corrélation presque parfaite" jusqu'à des longueurs de 100. Si toutes les séquences existantes pour  $N \leq 100$  respectent les structures utilisées, J. Wolfmann n'a pu démontrer que les séquences à corrélation quasi parfaites devaient respecter la structure qu'il avait remarquée sur des résultats expérimentaux ni qu'au-delà de la longueur 100 il n'existait d'autres séquences possédant les mêmes fonctions de corrélation.

L'étude présentée démontre que seules les séquences W.G. (Wolfmann-Goutelard) possèdent les fonctions de corrélation souhaitées et ont une structure unique dans laquelle entrent les "séquences à auto corrélation presque parfaite". Les propriétés particulières de ces séquences permettent de définir des algorithmes de construction rapide ainsi que des conditions d'existence ou de non existence pour certaines longueurs.

L'introduction de fonctions spécifiques utiles pour le développement est faite dans le chapitre 3.

## II. - RESULTATS ACQUIS ET DEFINITIONS -

Nous utiliserons comme base de départ les résultats acquis suivants :

- Les suites W.G. ne peuvent exister que dans des longueurs  $N$  multiples de 4.
- La fonction de corrélation d'une suite de symboles réels  $z_i$

$$(z_0, z_1, \dots, z_{N-1})$$

prend, pour un décalage  $k$  donné, la valeur

$$c_k = \sum_{i=0}^{N-1} z_i z_{i+k}$$

les indices étant définis modulo  $N$ .

Pour une suite binaire on donne habituellement les valeurs

$$\begin{aligned} z_i &= 1 && \text{si le symbole binaire est égal à } 0 \\ z_i &= -1 && \text{si le symbole binaire est égal à } 1. \end{aligned}$$

Avec cette correspondance, pour les séquences presque parfaites :

$$\begin{aligned} c_k &\equiv 0 \text{ excepté pour } c_0 = N \\ c_{N/2} &= -(N-4) \end{aligned}$$

## III. - TRANSFORMEES ET CORRELATIONS -

### III.1. Définitions

$$\begin{aligned} \text{Soit } F_2 &= \{0, 1\} \text{ le corps fini d'ordre } 2 \text{ et} \\ F_2^N &= \{g = (g_0, g_1, \dots, g_i, \dots, g_{N-1}) \mid g_i \in F_2\} \end{aligned} \quad (1)$$

On définit le vecteur  $a$  déduit de  $g$  par  
 $a = (a_0, a_1, \dots, a_i, \dots, a_{N-1})$  où  $a_i = (-1)^{g_i}$

On peut définir les suites périodiques  $g$  et  $a$  par les polynômes formels

$$g(x) = \sum_{i=0}^{N-1} g_i x^i \quad \text{dans } G(N) = F_2(x) / (x^N - 1) \quad (2)$$

$$\text{et } a(x) = \sum_{i=0}^{N-1} a_i x^i \quad \text{dans } R(N) = F_2(x) / (x^N - 1) \quad (3)$$

### III.2. Fonctions de corrélation

Soit deux séquences périodiques  $z_1(x)$  et  $z_2(x)$  définies par les polynômes formels

$$z_1(x) = \sum_{i=1}^{N-1} z_{1,i} x^i \quad \text{et} \quad z_2(x) = \sum_{i=1}^{N-1} z_{2,i} x^i$$

où  $z_{1,i}, z_{2,i} \in C$ .

On définit la fonction de corrélation périodique par la séquence

$$C_{1,2} = (c_0, c_1, \dots, c_k, \dots, c_{N-1}) \text{ telle que}$$

$$c_k = \sum_{i=0}^{N-1} z_{1,i}^* z_{2,i+k} = \sum_{i=0}^{N-1} z_{1,i-k}^* z_{2,i} \quad (4)$$

où les indices sont définis modulo  $N$  et où  $z^*$  dénote le conjugué de  $z$ .

La séquence  $C_{1,2}$  peut être représentée par le polynôme formel

$$C_{1,2}(x) = \sum_{k=0}^{N-1} c_k x^k \quad (5)$$

$C_{1,2}(x)$  s'exprime par le produit

$$C_{1,2}(x) = z_{1,1}(x) \cdot z_{2,1}(x) \quad (6)$$

Si  $z_2(x) \equiv z_1(x) = z(x)$  alors  $C_{1,2}(x) = C(x)$  est l'auto corrélation périodique.

Dans le cas de séquences réelles périodiques

$$r_1(x) = \sum_{i=0}^{N-1} r_{1,i} x^i \quad \text{et} \quad r_2(x) = \sum_{i=0}^{N-1} r_{2,i} x^i$$

où  $r_{1,i}, r_{2,i} \in R$  les relations précédentes se transposent sans difficulté et

$$C_{12}(x) = r_1(x) r_{2,1}(x)$$

On montre classiquement que le coefficient  $c_k$  se calcule à l'aide de la distance de Hamming  $d_{12}(k)$  entre la suite  $a_1(x)$  et la suite  $a_2(x) x^{-k}$  par la relation :

$$c_k = N - 2 d_{12}(k)$$

### III.3. $p$ transformée et $p$ permutée d'une séquence périodique

Dans la suite de cet exposé, nous noterons pour alléger la notation :

$$\sum_i \text{ la somme } \sum_{i=0}^{N-1} \text{ pour toute valeur entière de } i$$

$$\sum_{i0} \text{ la somme } \sum_{i=0}^{N-1} \text{ pour toutes les valeurs paires de } i$$

$$\sum_{i1} \text{ la somme } \sum_{i=0}^{N-1} \text{ pour toutes les valeurs impaires de } i$$

Les indices seront définis modulo N.

Soit une suite périodique  $z_{1(x)} = z(x) = \sum_i z_i x^i$

que nous appellerons séquence de base ou séquence un permutee.

On appellera séquence p permutee la séquence :

$$z_p(x) = z_{(xp)} = \sum_i z_i x^{pi} \quad \text{où } p \nmid N \quad \text{et } pp' = 1 \quad (7)$$

La séquence  $z_p(x)$  est alors unique et :

$$z_p(x) = \sum_i z_{ip'} x^i$$

Elle est obtenue par une permutation des éléments de la séquence  $z_1(x)$ .

On appellera p transformée d'une séquence  $z(x)$  la séquence définie par le polynôme

$$\tilde{z}_p(x) = \sum_h z(W_p^{-h}) x^h = \sum_h \tilde{z}_h x^h \quad (8)$$

$$\text{où } W_p = \exp(j \frac{2\pi}{N} p) \quad (9)$$

Il est simple de montrer que :

$z(x)$  et  $\tilde{z}_p(x)$  sont définies de façon bijective par la p transformée inverse

$$z(x) = \frac{1}{N} \sum_l \tilde{z}_p(W_p^l) x^l \quad (10)$$

#### III.4. p transformée d'une séquence réelle $r(x)$

D'après (12) alors

$$\tilde{r}_p(x) = \sum_h r(W_p^{-h}) x^h = \sum_h \tilde{r}_h x^h \quad (11)$$

Il est alors facile de démontrer que :

$$\text{- Si une séquence est réelle, alors } r_{(1/x)} = r(x)^* \quad (12)$$

$$\text{et alors } \tilde{r}_{N-h} = r(W_p^h) = \tilde{r}_h^* \quad (13)$$

- Si une séquence réelle est symétrique, alors  $\tilde{r}_h$  est réel et réciproquement.

Une séquence est symétrique si  $r_{(1/x)} = r(x)$  ce qui entraîne la double implication :

$$r(x) = r_{(1/x)} \iff \tilde{r}_p(x) = \tilde{r}_p(1/x) = \tilde{r}_p(x)^* \quad (14)$$

- Si une séquence est antisymétrique, alors  $\tilde{r}_h$  est imaginaire et réciproquement.

Une séquence est antisymétrique si  $r_{(1/x)} = -r(x)$  ce qui entraîne la double implication :

$$r(x) = -r_{(1/x)} \iff \tilde{r}_p(x) = -\tilde{r}_p(1/x) = -\tilde{r}_p(x)^* \quad (15)$$

#### III.5. La q transformée $z_p(x)$ d'une séquence $z_p(x)$ est la q' permutee de la un transformée de $z_p(x)$

Cette démonstration est évidente.

#### III.6. La q transformée d'une séquence p permutee $z_p(x)$ est la p transformée de la séquence q permutee $z_{pq}(x)$

La q transformée d'une séquence p permutee est donnée par

$$\tilde{z}_{pq}(x) = \sum_l z_p(W_q^{-l}) x^l \quad (16)$$

d'où l'on déduit sans peine :

$$\tilde{z}_{pq}(x) = \tilde{z}_{qp}(x) \quad (17)$$

#### III.7. p transformée des fonctions de corrélation périodiques

Soient les 2 séquences périodiques quelconques  $b_1(x)$  et  $b_2(x)$  définies précédemment et dont la fonction de corrélation est donnée par (6).

La p transformée vaut alors

$$\tilde{C}_{p12}(x) = \sum_h z_1^* (W_p^{-h}) z_2 (W_p^h) x^h \quad (18)$$

et dans le cas de l'autocorrélation

$$\tilde{C}_p(x) = \sum_h z^* (W_p^{-h}) z (W_p^h) x^h$$

$$\text{soit } \tilde{C}_p(x) = \sum_h |z(W_p^h)|^2 x^h \quad (19)$$

Dans le cas de suites réelles,

$$\tilde{C}_{p12}(x) = \sum_h r_1 (W_p^{-h}) r_2 (W_p^h) x^h$$

Si on pose  $r_1(W_p^{-h}) = A_1 W_0^{\alpha_1}$   
 $r_2(W_p^h) = A_2 W_0^{\alpha_2}$  (20)  
 avec  $W_0 = \exp(j\frac{\pi}{2})$

Alors  $C_{12}(x)$  est une fonction symétrique si et seulement si, d'après (14)

$\tilde{C}_{p12}(x)$  est à coefficients réels soit

$$\alpha_1 = \alpha_2 + 2\alpha \quad \text{où } \alpha \in \mathbb{Z}$$

$C_{12}(x)$  est une fonction antisymétrique si et seulement si, d'après (15)

$\tilde{C}_{p12}(x)$  est à coefficients imaginaires soit

$$\alpha_1 = \alpha_2 + 1 + 2\alpha \quad (21)$$

Il est important de remarquer que

si

$$C_{12}(x) = -C_{12}(1/x) \text{ et } r_2(x) = r_2(1/x) = > \\ r_1(x) = -r_1(1/x) \quad (22)$$

Soit  $r_2(x) = r_2(1/x)$

$$\text{alors de (12)} \quad \tilde{r}_{2p}(x) = \tilde{r}_{2p}(1/x) = \tilde{r}_{2p}^*(x)$$

et d'après (20)  $\alpha_2 = 0$ .

Alors d'après (21), (20) et (12)

$$\alpha_1 = 1 + 2\alpha, \tilde{r}_{1p}(x) = \tilde{r}_{1p}^* p^{-h} \text{ donc } \tilde{r}_{1p}(x) = -\tilde{r}_{1p}(1/x)$$

et d'après (15)  $r_1(x) = -r_1(1/x)$

#### IV. - STRUCTURES DES SEQUENCES W.G. -

##### IV.1. Définitions

Nous définirons les séquences W.G. par

$$g_1(x) = \sum_i g_i x^i \quad \text{ou } b_1(x) = \sum_i b_i x^i \\ \text{avec } g_i \in F_2 \quad \text{et } b_i = (-1)^{\delta_i}$$

Il est simple de montrer de  $C_{(N/2)} = -(N-4)$  implique que  $b_{N/2+i} = -b_i$  [7] excepté pour une valeur  $i_n$  de  $i$  pour laquelle on définit les éléments neutres  $b_{in} = b_{N/2+in}$ .

Il est donc toujours possible de choisir pour les séquences W.G. la forme canonique suivante :

$$b_1(x) = \sum_i b_i x^i$$

avec les éléments neutres  $b_{N/4} = b_{3N/4} = -1$

la condition initiale  $b_0 = 1$ ,  $b_{N/2} = -1$

la relation  $b_{N/2+i} = -b_i \quad \forall i \neq N/4, 3N/4$  (23)

et à laquelle correspondra

$$g_1(x) = \sum_i g_i x^i$$

##### IV.2. p transformée des séquences W.G.

Il est intéressant de poser :

$$b_1(x) = n(x) + b(x) \quad (24)$$

avec  $n(x) = x^{N/4} + x^{3N/4}$  composée des seuls éléments neutres

$$b(x) = \sum_{i \neq N/4 \text{ et } 3N/4} b_i x^i$$

La p transformée de  $n(x)$  est donc

$$\tilde{n}_p(x) = \sum_h (W_p^{-hN/4} + W_p^{-h3N/4}) x^h$$

compte tenu du fait que p est impair et en posant

$$\epsilon_h = [1 - (-1)^h] / 2$$

$$\tilde{n}_p(x) = \sum_h 2(1 - \epsilon_h) W_p^{-hN/4} x^h = \sum_h \tilde{n}_h x^h \quad (25)$$

Tous les coefficients des termes de puissance impaire sont nuls.

Tous les coefficients des termes de puissance paire ont un module égal à 2.

La p transformée de  $b(x)$  est donc :

$$\tilde{b}_p(x) = \sum_h b(W_p^{-h}) x^h$$

et s'écrit après développement :

$$\tilde{b}_p(x) = \sum_h \epsilon_h \left[ 2b_0 + 2 \sum_{i=1}^{\frac{N}{2}-1} (b_i W^{-hi} + b_{-i} W^{hi}) \right] x^h \quad (26)$$

La p transformée de  $b_1(x)$  est donnée avec (25) et (26) par

$$\tilde{b}_{1p}(x) = \tilde{n}_p(x) + \tilde{b}_p(x)$$

##### IV.3. q transformée de la fonction de corrélation des suites W.G.

La q transformée de la fonction de corrélation  $C(x)$  (31) est donnée par

$$\tilde{C}_q(x) = \sum_h [N - (N-4) W_q^{\frac{N}{2}h}] x^h$$

soit puisque  $W_q^{\frac{N}{2}} = (-1)$

$$\tilde{C}_q(x) = \sum_h [(N - (N-4) (-1)^h)] x^h = \sum_h \tilde{c}_h x^h$$

La q transformée est invariante avec q.

##### IV.4. Contraintes des p transformées des séquences W.G.

De la relation (18) appliquée au cas des séquences réelles il vient :

$$\tilde{c}_h = |\tilde{n}_h + \tilde{b}_h|^2$$

Soit :

- Si  $h$  est pair  $\tilde{c}_h = \left| \tilde{n}_h \right|^2$

relation toujours vérifiée  $\tilde{c}_h = 4, \left| \tilde{n}_h \right| = 2$

- Si  $h$  est impair  $\tilde{c}_h = 2N - 4 = \left| b_h \right|^2$

soit d'après (26)

$$2b_0 + 2 \sum_{i=1}^{\frac{N}{2}-1} (b_i W_p^{-hi} + b_{-i} W_p^{hi}) = A_0 W_0^{\alpha_0} \quad (27)$$

où  $A_0 = [2N - 4]^{1/2}$  et  $\alpha_0$  une fonction réelle de  $h$ .

#### IV.5. Structures de la sous séquence $b(x)$

Les contraintes sont imposées sur la séquence  $b(x)$  dont la  $p$  transformée ne comprend que des termes de puissance  $p$  impair. Nous retiendrons ce résultat pour la suite.

On peut écrire :

$$b(x) = p(x) + q(x)$$

où

$$p(x) = \sum_{i=0} b_i x^i \quad (i \text{ pair}) \quad (28)$$

$$q(x) = \sum_{i=1} b_i x^i \quad (i \text{ impair})$$

et pour lesquelles les coefficients des  $p$  transformées sont

$$(29) \quad \begin{aligned} \tilde{p}_h &= \sum_{i=0} b_i W_p^{-hi} \quad \text{et} \quad \tilde{p}_{\frac{N}{2}+h} = \sum_{i=0} b_i W_p^{-(\frac{N}{2}+h)i} = \tilde{p}_h \\ \tilde{q}_h &= \sum_{i=1} b_i W_p^{-hi} \quad \text{et} \quad \tilde{q}_{\frac{N}{2}+h} = \sum_{i=1} b_i W_p^{-(\frac{N}{2}+h)i} = -\tilde{q}_h \end{aligned}$$

puisque  $W_p^{\frac{N}{2}} = (-1)$

de (27) (28) (29) il vient donc

$$\left| \tilde{p}_h + \tilde{q}_h \right| = \left| \tilde{p}_h - \tilde{q}_h \right|$$

soit en posant  $\tilde{p}_h = A_p W_0^{\alpha_p}$

et  $\tilde{q}_h = A_q W_0^{\alpha_q}$

$$\left| A_p + A_q W_0^{(\alpha_q - \alpha_p)} \right| = \left| A_p - A_q W_0^{(\alpha_q - \alpha_p)} \right|$$

Cette relation est vérifiée si, et seulement si :

$$\alpha_q = \alpha_p + 1 + 2\alpha$$

D'après (21) il s'ensuit que la fonction de corrélation périodique  $C_{pq}(x)$  entre les séquences  $p(x)$  et  $q(x)$  est antisymétrique.

Les séquences  $p(x)$  et  $q(x)$  sont toujours décomposables de la façon suivante :

$$\begin{aligned} p(x) &= p'(x) + p''(x) \\ q(x) &= q'(x) + q''(x) \end{aligned} \quad (30)$$

ou  $p'(x) = p'(1/x), q'(x) = q'(1/x)$   
sont des séquences paires

$$p''(x) = -p''(1/x), q''(x) = -q''(1/x)$$

sont des séquences impaires

Si on pose

$$p'(x) = \sum_{i=0} p'_i x^i : p''(x) = \sum_{i=0} p''_i x^i$$

$$q'(x) = \sum_{i=1} q'_i x^i : q''(x) = \sum_{i=1} q''_i x^i$$

alors

$$p'_i = p''_i = 0 \quad \text{pour } i = \frac{N}{4} \text{ impair si } N \equiv 0 \pmod{8}$$

$$q'_i = q''_i = 0 \quad \text{pour } i = \frac{N}{4} \text{ pair si } N \not\equiv 0 \pmod{8}$$

$$p'_1 = p'_{N-1} (q'_1 = q'_{N-1}) = b_1 \text{ et}$$

$$p''_1 = p''_{N-1} (q''_1 = q''_{N-1}) = 0 \text{ si } b_{N-1} = b_1 \neq 0$$

$$p''_1 = -p''_{N-1} (q''_1 = -q''_{N-1}) = b_1 \text{ et}$$

$$p'_1 = p'_{N-1} (q'_1 = q'_{N-1}) = 0 \text{ si } b_{N-1} = -b_1 \neq 0$$

La fonction de corrélation s'écrit alors

$$C_{pq}(x) = C_{p'q'}(x) + C_{p''q''}(x) + C_{p'q''}(x) + C_{p''q'}(x)$$

Les deux dernières fonctions sont antisymétriques. En effet :

$$C_{p'q''}(x) = p'(x) q''(1/x) = -p'(1/x) q''(x) = -C_{p''q'}(1/x)$$

Les deux premières fonctions sont symétriques.

$$C_{p'q'}(x) = p'(x) q'(1/x) = p'(1/x) q'(x) = C_{p'q'}(1/x)$$

$$C_{p''q''}(x) = p''(x) q''(1/x) = p''(1/x) q''(x) = C_{p''q''}(1/x)$$

La fonction de corrélation est donc antisymétrique si et seulement si

$$C_{p'q'}(x) + C_{p''q''}(x) = 0 \quad (31)$$

Si l'on pose

$$C_{p'q'}(x) = \sum_k c'_k x^k \text{ et } C_{p''q''}(x) = \sum_k c''_k x^k$$

la condition (31) impose

$$c'_k + c''_k = 0 \quad \forall k \quad (32)$$

relation toujours vérifiée pour  $k$  pair et pour  $k$  impair, d'après (4)

$$c'_k = \sum_{i=0} p'_i q'_{i+k} \text{ et } c''_k = \sum_{i=0} p''_i q''_{i+k}$$

Si la relation (32) est vérifiée,  $1/2C'_k$  et  $1/2C''_k$  ont évidemment la même parité.

Si on considère les suites  $b(x)$ ,  $p(x)$  et  $q(x)$ , d'après (24) et (28)

$b(x)$  comporte  $N - 2$  termes non nuls

(33)

$p(x)$  comporte  $P$  termes non nuls :

$$P = \frac{N}{2} - 2 \quad \text{si } N \equiv 0 \pmod{8}$$

car le degré des éléments neutres est pair

$$P = \frac{N}{2} \quad \text{si } N \not\equiv 0 \pmod{8}$$

$q(x)$  comporte  $Q$  termes non nuls :

$$Q = \frac{N}{2} \quad \text{si } N \equiv 0 \pmod{8}$$

$$Q = \frac{N}{2} - 2 \quad \text{si } N \not\equiv 0 \pmod{8}$$

$\frac{P}{2}$  est donc toujours impair,  $\frac{Q}{2}$  est toujours pair.

On posera  $P = P' + P''$  où  $P'$  et  $P''$  sont

respectivement le nombre de termes non nuls de  $p'(x)$

et  $p''(x)$ .

De même  $Q = Q' + Q''$  où  $Q'$  et  $Q''$  sont

respectivement le nombre de termes non nuls des

séquences  $q'(x)$  et  $q''(x)$ .

Si on considère  $c'_k$ , on peut distinguer 3 types de termes

-  $P'_k$  termes non nuls -

$P'_k$  est pair car  $p'_i q'_{i+k} = p'_{i+\frac{N}{2}} q'_{i+k+\frac{N}{2}}$

Ces termes n'ayant que deux valeurs possibles  $+1$  ou  $-1$

$P'_k = P'_{k+} + P'_{k-}$  où  $P'_{k+}$  et  $P'_{k-}$  sont

respectivement les nombres de termes prenant la valeur  $+1$  et  $-1$ .

Donc  $c'_k = P'_{k+} - 2P'_{k-}$  et comme  $P'_{k-}$  est pair  $c'_k/2$  à la parité de  $P'_{k+}/2$

-  $Z_k$  termes nuls par le fait qu'un seul des termes  $p'_i$  ou  $q'_{i+k} = 0$  et pour  $i \neq \frac{N}{4}$  et  $i+k \neq \frac{N}{4}$  (34)

- Les termes nuls situés à la place des éléments neutres

(35)

. apparaissent dans  $p' \frac{N}{4}$  et  $p' \frac{3N}{4}$  si  $N \equiv 0 \pmod{8}$  et

si  $q' \frac{N}{4} + k \neq 0$  on posera  $\epsilon_k = 1$  sinon on posera  $\epsilon_k = 0$

. apparaissent dans  $q' \frac{N}{4}$  et  $q' \frac{3N}{4}$  si  $N \not\equiv 0 \pmod{8}$  et

si  $p' \frac{N}{4} - k \neq 0$  on posera  $\epsilon_k = 1$  sinon on posera  $\epsilon_k = 0$

Le nombre total de termes des deux séquences  $p'(x)$  et  $q'(x)$  vaut donc

$$P' + Q' = 2P'_k + Z_k + 2\epsilon_k$$

Pour le terme  $c''_k$  on distingue les mêmes types de termes

-  $Q'_k$  termes non nuls et pour les mêmes raisons que précédemment

$$c''_k \text{ à la parité de } Q'_k/2 \quad (36)$$

-  $Z_k$  termes nuls par le fait qu'un seul des termes  $p''_i$  ou  $q''_{i+k} = 0$  et pour  $i \neq \frac{N}{4}$  et  $i+k \neq \frac{N}{4}$ .

En effet d'après (28) (30) et (34)

$$p'_i = 0 \text{ et } q'_{i+k} \neq 0 \Leftrightarrow p''_i \neq 0 \text{ et } q''_{i+k} = 0$$

$$p'_i \neq 0 \text{ et } q'_{i+k} = 0 \Leftrightarrow p''_i = 0 \text{ et } q''_{i+k} \neq 0$$

- Les termes nuls situés à la place des éléments neutres :

(37)

. apparaissent dans  $p'' \frac{N}{4}$  et  $p'' \frac{3N}{4}$  si  $N \equiv 0 \pmod{8}$  et

d'après (28) (30) et (35)

$$q'' \frac{N}{4} + k \neq 0 \text{ si } \epsilon_k = 0$$

. apparaissent dans  $q'' \frac{N}{4}$  et  $q'' \frac{3N}{4}$  si  $N \not\equiv 0 \pmod{8}$  et

d'après (28) (30) et (35)

$$p'' \frac{N}{4} - k \neq 0 \text{ si } \epsilon_k = 0$$

Le nombre total de termes des deux séquences  $p''(x)$  et  $q''(x)$  vaut donc :

$$P'' + Q'' = 2Q'_k + Z_k + 2(1 - \epsilon_k)$$

$c'_k/2$  et  $c''_k/2$  ayant même parité, (35), (38) conduisent à

$$\frac{P' + Q' - P' - Q' - 2}{4} + \epsilon_k \equiv 0 \pmod{2} \quad (38)$$

Cette relation montre que  $\epsilon_k$  doit être invariant avec

$k$ , ce qui implique que dans chaque doublet

$\{p'(x), q'(x)\}$  et  $\{p''(x), q''(x)\}$  la séquence qui ne comporte pas les éléments neutres doit être, soit nulle,

soit complète, c'est-à-dire que tous les coefficients des valeurs possibles de  $i(i_0$  ou  $i_1)$  soient différents de zéro.

Cette séquence est donc, soit symétrique, soit

antisymétrique. D'après (22) l'autre séquence est donc

respectivement antisymétrique ou symétrique.

Les seules configurations possibles sont donc :

$$p(x) = p(1/x) \quad \text{et} \quad q(x) = -q(1/x)$$

$$\text{ou} \quad p(x) = -p(1/x) \quad \text{et} \quad q(x) = q(1/x)$$

Cette dernière solution est impossible car  $p'(x) \neq 0$  puisqu'elle contient  $b_0 = 1$

La seule solution possible est donc

$$p''(x) = 0 \quad \text{et} \quad q'(x) = 0 \quad (39)$$

La relation (38) s'écrit alors

$$\frac{Q - P - 2}{4} + \epsilon_k \equiv \text{mod } 2$$

soit, compte tenu de la relation (33) en posant

$$\epsilon_N = 0 \text{ si } N \equiv 0 \pmod{8}$$

$$\epsilon_N = 1 \text{ si } N \not\equiv 0 \pmod{8}$$

$$P = \frac{N}{2} - 2(1 - \epsilon_N) \quad Q = \frac{N}{2} - 2 \in \mathbb{N}$$

ce qui entraîne

$$\epsilon_k = \epsilon_N$$

#### IV.6. Structure des séquences W.G.

Les séquences W.G. sont donc définies

- par leur forme canonique (23)

$$b_1(x) = \sum_i b_i x^i$$

avec les éléments neutres  $b_{\frac{N}{2}} = b_{\frac{3N}{4}} = 1$

Les conditions initiales  $b_0 = 1 \quad b_{\frac{N}{2}} = -1$

$$b_{\frac{N}{2}+i} = -b_i \quad \forall i \neq \frac{N}{4}, \frac{3N}{4}$$

- De (24), (28) et (39)

$$b_1(x) = n(x) + p(x) + q(x)$$

avec

$$n(x) = x^{\frac{N}{4}} + x^{\frac{3N}{4}}$$

$$p(x) = \sum_{i \text{ pair}} b_i x^i$$

$$q(x) = \sum_{i \text{ impair}} b_i x^i$$

$$\tilde{n}_p(x) = \sum_h \tilde{n}_h x^h, \quad h \text{ pair}, \quad |\tilde{n}_h| = 2$$

$$\tilde{p}_p(x) = \sum_h \tilde{p}_h x^h, \quad h \text{ impair}, \quad \tilde{p}_h \text{ réel}$$

$$\tilde{q}_p(x) = \sum_h \tilde{q}_h x^h, \quad h \text{ impair}, \quad \tilde{q}_h = \text{imaginaire}$$

et la relation :

$$|\tilde{p}_h|^2 + |\tilde{q}_h|^2 = 2N - 4 \quad (40)$$

- Les propriétés des séquences  $n(x)$ ,  $p(x)$ ,  $q(x)$  sont

(41)

$$n(xP) = n(x), \quad n(x^{\frac{N}{2}+P}) = n(x), \quad n(x^{N-P}) = n(x),$$

$$x^{\frac{N}{2}} n(xP) = n(x)$$

$$p(x \cdot P) = p(xP), \quad x^{\frac{N}{2}} p(xP) = -p(xP),$$

$$p(x^{\frac{N}{2}+P}) = p(xP) = p(x \cdot P)$$

$$q(x \cdot P) = -q(xP), \quad x^{\frac{N}{2}} q(xP) = -q(xP),$$

$$q(x^{\frac{N}{2}+P}) = -q(xP) = q(x \cdot P)$$

- Les propriétés déduites de (39)

$$b_{N-i} = (-1)^i b_i \quad i \neq \frac{N}{4}, \frac{3N}{4}$$

On associe à cette définition (42)

$$g_1(x) = \sum g_i x^i$$

$$\text{où } b_i = (-1)^{g_i}$$

$b_1(x)$  et  $g_1(x)$  s'expriment donc par :

(43)

$$b_1 = (1 \quad b_1 \quad b_2 \dots b_{\frac{N}{4}-1}, \quad b_{\frac{N}{4}} = 1, \quad b_{\frac{N}{4}+1} = (-1)^{\epsilon_N} b_{\frac{N}{4}-1}, \dots, -b_2, b_1)$$

$$-1 \quad -b_1 \quad -b_2 \dots -b_{\frac{N}{4}-1}, \quad b_{\frac{3N}{4}} = 1, \quad (-1)^{\epsilon_N+1} b_{\frac{N}{4}-1} \dots$$

$$\dots b_2, -b_1)$$

et

$$g_1 = (0 \quad g_1 \quad g_2 \dots g_{\frac{N}{4}-1}, \quad g_{\frac{N}{4}} = 0, \quad g_{\frac{N}{4}+1} = g_{\frac{N}{4}-1} + \epsilon_N + 1, \dots, \bar{g}_2, g_1)$$

$$1 \quad g_1 \quad g_2 \dots g_{\frac{N}{4}-1}, \quad g_{\frac{3N}{4}} = 0, \quad g_{\frac{3N}{4}+1} = g_{\frac{N}{4}-1} + \epsilon_N, \dots, g_2, \bar{g}_1)$$

où  $\bar{g}_i$  dénote le complément dans  $F_2$  de  $g_i$ .  $g_i + \bar{g}_i = 1$

La forme canonique de  $g_1(x)$  est de la forme

$$g_1 = (0, G, 0, G^0, 1, \bar{G}, 0, \bar{G}^0)$$

avec  $G = (g_1, g_2 \dots g_{\frac{N}{4}-1})$

$$G^0 = (g_{\frac{N}{4}-1} + \epsilon_N, g_{\frac{N}{4}-1} + \epsilon_N+1, \dots, \bar{g}_2, g_1)$$

qui est la séquence  $G$  inversée dont les éléments d'ordre impair sont complémentés.

$\bar{G}$  et  $\bar{G}^0$  sont les séquences complémentées de  $G$  et  $G^0$ .

#### IV.7. Conclusion sur l'existence et la structure des séquences W.G.

L'étude développée dans ce paragraphe montre que toutes les séquences W.G. respectent les conditions précédentes et qu'il n'existe pas d'autres types de séquences.

Les seules structures possibles de ces séquences et leurs propriétés générales sont décrites par les relations (39) à (43).

La relation (43) qui donne la forme canonique des séquences W.G. montre qu'elles sont définies par  $N/4 - 1$  symboles.

#### V. - PROPRIETES DES SEQUENCES -

Les propriétés des séquences W.G. se déduisent directement des relations (39) à (43).

##### V.1. Propriété 1

Toutes les permutations circulaires conservent les propriétés de la corrélation périodiques des séquences W.G.

Cette propriété est propre à toutes les séquences périodiques.

##### V.2. Propriété 2

Le poids des séquences W.G. est égal à 2. Le poids de la partie formée par les termes pairs est égal à  $2\epsilon_N$  et celui de la partie formée par les termes impairs est égal à  $2(1 - \epsilon_N)$ .

En effet, d'après (29)

$$\tilde{p}_h=0 = \left| \sum_{i_0} b_i \right| = w_p$$

$$\tilde{q}_h=0 = \left| \sum_{i_1} b_i \right| = w_q$$

où  $w_p$  et  $w_q$  sont les poids de  $p(x)$  et  $q(x)$ .

Comme  $b_{\frac{N}{2}+i} = -b_i$

$$w_p = w_q = 0$$

Si  $N \equiv 0 \pmod{8}$  les termes neutres sont dans la suite paire et dans le cas contraire dans la suite impaire, ce qui démontre la proposition.

##### V.3. Propriété 3

Les  $p$  permutées  $g_p(x) = g_1(x^p)$  de la séquence de base  $g_1(x)$  sont également des séquences W.G.

En effet, on a montré que quel que soit  $p$ , la  $p$  transformée de la séquence de base satisfaisait les conditions de corrélation.

On a montré que la  $q$  transformée d'une séquence  $p$  permutée est la  $p$  transformée d'une séquence  $q$  permutée.

La  $p$  transformée de la suite de base est donc la 1 transformée de sa  $p$  permutée, qui satisfait donc les conditions de corrélation.

##### V.4. Propriété 4

Les séquences complémentées sont des translatés de séquences à éléments neutres inversés.

Une séquence  $p$  permutée complémentée est définie par

$$\bar{g}_1(x^p) = \sum \bar{g}_1 x^{ip}$$

Il correspond la suite  $\bar{b}_1(x)$

$$\bar{b}_1(x^p) = -n(x^p) - p(x^p) - q(x^p)$$

soit d'après (41)

$$b_1(x^p) = [-n(x^p) + p(x^p) + q(x^p)] x^{\frac{N}{2}}$$

ou

$$b_1(x^p) = x^{\frac{N}{2}} \overset{c}{b}_1(x^p) \quad (45)$$

dans laquelle

$\overset{c}{b}_1(x^p)$  se déduit de  $b_1(x^p)$  en complémentant les éléments neutres.

##### V.5. Propriété 5

Les séquences alternées sont des séquences W.G.

On appellera

- Séquences alternées impaires  $g_a(x^p)$  celles dont la valeur des éléments impairs de  $g_1(x^p)$  est complémentée

$$g_a(x^p) = g_1(x^p) + I(x)$$

$$\text{avec } I(x) = \sum_{i_1} x^i$$

- Les séquences alternées paires sont les séquences alternées impaires complémentées

$$\bar{g}_a(x^p) = g_1(x^p) + \bar{I}(x) = g_1(x^p) + P(x)$$

$$\text{avec } P(x) = \sum_{i_0} x^i$$

Alors

$$b_a(x^p) = (-1)^{\epsilon_N} n(x^p) + p(x^p) - q(x^p) \quad (46)$$



qui donne d'après (40)

$|\tilde{n}_h| = 2 \quad \tilde{p}_h^2 + |\tilde{q}_h|^2 = 2N - 4$   
qui vérifient les conditions spectrales des séquences W.G.

$b_a(x)$  et son complément  $\bar{b}_a(x)$  sont donc des séquences W.G.

#### V.6. Propriété 6

Les séquences inverses  $g_1(x^{-p})$  des séquences  $g_1(xP)$  sont des séquences alternées de  $g_1(xP)$ . Elles sont à éléments neutres inversés si  $N \not\equiv 0 \pmod{8}$ .

Si  $p \nmid N$  alors  $N - p \nmid N$

et  $b_1(x^{-p}) = n(x^{-p}) + p(x^{-p}) + q(x^{-p})$

D'après (41)  $b_1(x^{-p}) = n(xP) + p(xP) - q(xP)$

Il s'ensuit en comparant avec (46) que

$$b_1(x^{-p}) = b_a(xP) \quad \text{si } \epsilon_N = 0 \text{ soit } N \equiv 0 \pmod{8}$$

$$b_1(x^{-p}) = b_a^c(xP) = x^{\frac{N}{2}} \overline{b_a(xP)} \quad \text{si } \epsilon_N = 1 \text{ soit } N \not\equiv 0 \pmod{8}$$

#### V.7. Propriété 7

Les séquences permutées  $g_1(xP)$  et  $g_1(x^{\frac{N}{2}-p})$  sont égales

si  $p \nmid N$  alors  $\frac{N}{2} - p \nmid N$

et  $b_1(xP) = n(xP) + p(xP) + q(xP)$

D'après (41)

$$n(x^{\frac{N}{2}-p}) = n(xP) \quad p(x^{\frac{N}{2}-p}) = p(xP) \quad q(x^{\frac{N}{2}-p}) = q(xP)$$

donc

$$g_1(xP) = g_1(x^{\frac{N}{2}-p})$$

#### V.8. Propriété 8

La  $p$  permutée de la translatée de la forme de base  $g_1(x)$  est une permutée de  $g_1(xP)$

soit  $g'_1(x) = x^\alpha g_1(x)$

$$g'_1(xP) = x^{\alpha P} g_1(xP)$$

#### V.9. Conséquences des propriétés des séquences W.G.

A partir de la séquence de base  $b_1(x)$  on peut constater que les séquences

$b_1(xP)$ ,  $b_1(x^{N-p})$ ,  $b_1(x^{\frac{N}{2}-p})$ ,  $b_1(x^{\frac{N}{2}+p})$  sont liées simplement :

Soit

$$b_1(x) = \sum_i b_i x^i \quad b_p(x) = b_1(xP) = \sum_i b_i x^{iP}$$

alors

$$b_1(x^{\frac{N}{2}-p}) = b_1(xP)$$

$$b_1(x^{\frac{N}{2}+p}) = b_1(x^{N-p})$$

$$b_1(x^{N-p}) = \begin{cases} b_a(xP) & \text{si } N \equiv 0 \pmod{8} \\ b_a^c(xP) & \text{si } N \not\equiv 0 \pmod{8} \end{cases}$$

En particulier si  $p = 1$

$$b_1(x^{\frac{N}{2}-1}) = b_1(x)$$

$$b_1(x^{\frac{N}{2}+1}) = b(x^{-1})$$

$$b_1(x^{N-1}) = \begin{cases} b_a(x) & \text{si } N \equiv 0 \pmod{8} \\ b_a^c(x) & \text{si } N \not\equiv 0 \pmod{8} \end{cases}$$

#### VI. - EXISTENCE ET CONSTRUCTION DES SEQUENCES W.G. -

La relation (27) doit être vérifiée quel que soit  $h$  impair.

Comme toute  $p$  permutée séquence W.G. est également une séquence W.G., il suffit que cette relation soit vérifiée pour  $p = 1$ .

La relation s'écrit alors :

$$2b_0 + 2 \sum_{i=1}^{\frac{N}{2}-1} [b_i W_1^{-hi} + b_{-i} W_1^{hi}] = [2N-4]^{1/2} W_0^{\alpha_0} \quad (47)$$

Les séquences peuvent être construites à partir de cette relation, compte tenu des propriétés des séquences W.G.

Si on appelle

$$P_0 = \# \{ \text{symboles égaux à } b_0 \text{ de } p(x) \}$$

$$P_1 = \# \{ \text{symboles égaux à } -b_0 \text{ de } p(x) \}$$

$$Q_0 = \# \{ \text{symboles égaux à } 1 \text{ de } q(x) \}$$

$$Q_1 = \# \{ \text{symboles égaux à } -1 \text{ de } q(x) \}$$

$$p = |P_0 - P_1|$$

$$q = |Q_0 - Q_1|$$

La relation (47) impose :

$$\begin{aligned} & \text{Pour } \frac{N}{4} \neq 0 \bmod 2 \\ & (1 + 2q)^2 + (2p)^2 = \frac{N}{2} - 1 \end{aligned} \quad (48)$$

$$\begin{aligned} & \text{Pour } \frac{N}{8} \neq 0 \bmod 2 \\ & (1 + 2p)^2 + (2q)^2 = \frac{N}{2} - 1 \end{aligned} \quad (49)$$

$$\begin{aligned} & \text{Pour } \frac{N}{16} \neq 0 \bmod 2 \\ & 2p_2(1 + 2p_1) + q_1^2 - q_2^2 + 2q_1q_2 = 0 \\ & \text{et} \\ & (1 + 2p_1)^2 + 2p_2^2 + 2(q_1^2 + q_2^2) = \frac{N}{2} - 1 \\ & \text{avec } p = p_1 + p_2 \\ & q = q_1 + q_2 \end{aligned} \quad (50)$$

Les relations (48), (49) et (50) permettent de déterminer les conditions de non existence des séquences W.G.

Les tableaux 1, 2, 3, donnent pour les longueurs inférieures à 20 000 celles pour lesquelles il n'existe pas de séquences W.G. Ces tableaux révèlent qu'il existe un grand nombre de séquences W.G. avec un choix de longueurs très riche.

Le tableau 4 donne un exemple de ces séquences pour les longueurs  $N/4 = 0 \bmod 2$  inférieures à 150.

On a porté les séquences de base desquelles on peut déduire :

- Les séquences  $q$  permutées dont le nombre est pour

$$N = 2^\alpha 3^\beta 5^\gamma \dots e^\epsilon \quad \text{donné}$$

par la relation

$$\Phi(N) = N \left(1 - \frac{1}{a}\right) \left(1 - \frac{1}{b}\right) \dots \left(1 - \frac{1}{e}\right)$$

- Les séquences complémentées obtenues en complétant les symboles de la séquence.

- Les séquences obtenues par permutations circulaires, par complément des symboles pair ou impair et par une écriture inversée des symboles sont contenues, comme il l'a été démontré, dans les séquences  $q$  permutées ou complémentées.

Chaque séquence de base permet la génération des  $2 \Phi(N)$  séquences de même longueur indiquées dans le tableau 4.

Ainsi pour :

$$\begin{aligned} N = 1004 & & 2 \Phi(1004) = 500 \\ N = 10004 & & 2 \Phi(10004) = 4800 \\ N = 20012 & & 2 \Phi(20012) = 10004 \end{aligned}$$

La croissance rapide du nombre de séquences existantes avec  $N$  est une propriété très intéressante dans les applications.

La figure 2 donne les séquences  $q$  permutées de longueur 28, auxquelles il faut ajouter les séquences complémentées, et sur lesquelles on peut retrouver toutes les propriétés énoncées précédemment.

## VII. - CONCLUSION -

Les séquences périodiques binaires W.G. sont les seules à présenter une fonction de corrélation nulle en dehors du pic central et d'un seul décalage. Ces fonctions de corrélation sont les meilleures que l'on puisse obtenir avec des séquences binaires.

Il a été montré que toute séquence présentant ces caractéristiques était une séquence W.G. et donc qu'il n'en existe pas d'autre.

Les structures de ces séquences ont été définies et les longueurs dans lesquelles il est possible de les construire sont des multiples de 4. Cette caractéristique est intéressante car elle permet d'offrir un grand choix aux utilisateurs. Leur non existence a été déterminée jusqu'aux longueurs 20 000 et des exemples de séquences W.G. sont donnés.

L'utilisation de ces séquences offre un intérêt certain dans les problèmes de radiolocalisation, de télédétection et de télécommunication où les techniques de codage offrent des solutions séduisantes.

Le LETTI a appliqué ces techniques au sondage ionosphérique par rétrodiffusion pour réduire la puissance d'émission utilisée. Il est connu que les puissances à mettre en jeu sont importantes à cause des pertes dues à la propagation. La figure 3 représente le progrès qui a été obtenu par l'utilisation des techniques de codage en deux décennies. En 1962, la technique impulsionnelle nécessitait une puissance d'émission de 100kW. En 1970, le LETTI a mis en oeuvre, dans un système de sondeur à rétrodiffusion monostatique, une première méthode de codage qui a permis de réduire la puissance d'émission à 1kW. Actuellement, ces techniques permettent, à qualités signal/bruit identiques, de réaliser ces mesures avec des puissances d'émission inférieures à 1W. Cette illustration révèle la puissance de ces techniques que les séquences W.G. peuvent contribuer à accroître.

- [1] R. SIVASWAMY - Multiphase complementary Codes. IEEE Trans. Infor. Theory, Vol IT-24, n° 5, sept. 1978.
- [2] C. GOUTELARD - Séquences d'étalement à synchronisation optimale par commutations d'opérateurs incommutables. Conf. AGARD-EPP, n° 442, Paris, oct. 1988.
- [3] R. HOHOLDT, J. JUSTEGEN - Ternary sequences with perfect periodic autocorrelation. IEEE Trans. Info. Theory, Vol. IT-29, pp 597-599, 1983.
- [4] L. D. BAUMERT - Cyclic difference sets. Lecture notes in mathematics, Vol. 182, Springer-Verlag, 1971.
- [5] C. GOUTELARD, F. CHAVAND - Codes pseudo-orthogonaux adaptés aux transmissions dans les canaux à caractéristiques aléatoires. Annales des Télécomm., n° 11-12, Tome 30, sept. 1975.
- [6] F.J. POLGE - A general solution for the synthesis of binary sequences with desired correlation sequence. AGARD Conf. 381 AVP Multifunction radar for airborne application, Toulouse. oct. 1985.
- [7] R. ALEXIS - Search for sequences with zero autocorrelation. Lecture notes in computer Science, Vol. 311, pp 159-172, Springer-Verlag, 1987.
- [8] J. WOLFMANN - Almost perfect autocorrelation sequences. Eurocode 90, Udine (Italie), nov. 1990.



14728	14712	14696	14680	14664	8664	8696	8712	8744	14728	14712	14696	14680	14664	8664	8696	8712	8744
14856	14792	14760	14744	14728	8672	8808	8872	8920	14856	14792	14760	14744	14728	8672	8808	8872	8920
15032	14968	14952	14936	14920	9112	9144	9192	9240	15032	14968	14952	14936	14920	9112	9144	9192	9240
15192	15160	15144	15128	15112	9256	9288	9336	9384	15192	15160	15144	15128	15112	9256	9288	9336	9384
15320	15304	15288	15272	15256	9432	9464	9512	9560	15320	15304	15288	15272	15256	9432	9464	9512	9560
15480	15464	15448	15432	15416	9592	9624	9672	9720	15480	15464	15448	15432	15416	9592	9624	9672	9720
15576	15560	15544	15528	15512	9736	9768	9816	9864	15576	15560	15544	15528	15512	9736	9768	9816	9864
15688	15672	15656	15640	15624	9912	9944	9992	10040	15688	15672	15656	15640	15624	9912	9944	9992	10040
15832	15816	15800	15784	15768	10072	10104	10152	10200	15832	15816	15800	15784	15768	10072	10104	10152	10200
15992	15976	15960	15944	15928	10248	10280	10328	10376	15992	15976	15960	15944	15928	10248	10280	10328	10376
16088	16072	16056	16040	16024	10392	10424	10472	10520	16088	16072	16056	16040	16024	10392	10424	10472	10520
16264	16248	16232	16216	16200	10600	10632	10680	10728	16264	16248	16232	16216	16200	10600	10632	10680	10728
16424	16408	16392	16376	16360	10712	10744	10792	10840	16424	16408	16392	16376	16360	10712	10744	10792	10840
16616	16600	16584	16568	16552	10904	10936	10984	11032	16616	16600	16584	16568	16552	10904	10936	10984	11032
16712	16696	16680	16664	16648	11040	11072	11120	11168	16712	16696	16680	16664	16648	11040	11072	11120	11168
16984	16968	16952	16936	16920	11160	11192	11240	11288	16984	16968	16952	16936	16920	11160	11192	11240	11288
17112	17096	17080	17064	17048	11272	11304	11352	11400	17112	17096	17080	17064	17048	11272	11304	11352	11400
17208	17192	17176	17160	17144	11432	11464	11512	11560	17208	17192	17176	17160	17144	11432	11464	11512	11560
17432	17416	17400	17384	17368	11544	11576	11624	11672	17432	17416	17400	17384	17368	11544	11576	11624	11672
17576	17560	17544	17528	17512	11752	11784	11832	11880	17576	17560	17544	17528	17512	11752	11784	11832	11880
17688	17672	17656	17640	17624	11912	11944	11992	12040	17688	17672	17656	17640	17624	11912	11944	11992	12040
17832	17816	17800	17784	17768	12072	12104	12152	12200	17832	17816	17800	17784	17768	12072	12104	12152	12200
17912	17896	17880	17864	17848	12232	12264	12312	12360	17912	17896	17880	17864	17848	12232	12264	12312	12360
18104	18088	18072	18056	18040	12376	12408	12456	12504	18104	18088	18072	18056	18040	12376	12408	12456	12504
18264	18248	18232	18216	18200	12504	12536	12584	12632	18264	18248	18232	18216	18200	12504	12536	12584	12632
18424	18408	18392	18376	18360	12664	12696	12744	12792	18424	18408	18392	18376	18360	12664	12696	12744	12792
18520	18504	18488	18472	18456	12872	12904	12952	13000	18520	18504	18488	18472	18456	12872	12904	12952	13000
18696	18680	18664	18648	18632	13112	13144	13192	13240	18696	18680	18664	18648	18632	13112	13144	13192	13240
18872	18856	18840	18824	18808	13272	13304	13352	13400	18872	18856	18840	18824	18808	13272	13304	13352	13400
19000	18984	18968	18952	18936	13464	13496	13544	13592	19000	18984	18968	18952	18936	13464	13496	13544	13592
19192	19176	19160	19144	19128	13672	13704	13752	13800	19192	19176	19160	19144	19128	13672	13704	13752	13800
19336	19320	19304	19288	19272	13848	13880	13928	13976	19336	19320	19304	19288	19272	13848	13880	13928	13976
19448	19432	19416	19400	19384	13992	14024	14072	14120	19448	19432	19416	19400	19384	13992	14024	14072	14120
19560	19544	19528	19512	19496	14136	14168	14216	14264	19560	19544	19528	19512	19496	14136	14168	14216	14264
19736	19720	19704	19688	19672	14344	14376	14424	14472	19736	19720	19704	19688	19672	14344	14376	14424	14472
19896	19880	19864	19848	19832	14584	14616	14664	14712	19896	19880	19864	19848	19832	14584	14616	14664	14712
20024	20008	19992	19976	19960	14824	14856	14904	14952	20024	20008	19992	19976	19960	14824	14856	14904	14952

TABLEAU 2

Longueur  $N/8 \neq 0 \bmod 2$  pour lesquelles les séquences W.G. n'existent pas. $N \leq 20000$ 

850 longueurs existantes

400 longueurs non existantes

80	112	176	272	368	432	464	496	560	592	656	752	816	848	912	94
1040	1072	1104	1168	1232	1328	1392	1424	1520	1552	1616	1712	1744			
1808	1872	1904	2000	2032	2096	2160	2192	2224	2288	2320	2352	2384			
2480	2512	2544	2576	2672	2768	2832	2864	2928	2960	2992	3056	3184			
3248	3280	3312	3344	3408	3440	3472	3536	3632	3792	3824	3856	3888			
3920	3952	3984	4016	4112	4144	4208	4240	4272	4304	4336	4368	4400			
4432	4496	4528	4560	4592	4656	4688	4752	4784	4816	4880	4912	4976			
5040	5072	5168	5232	5264	5360	5392	5456	5488	5520	5552	5648	5712			
5744	5840	5872	5904	5936	5968	6032	6096	6128	6192	6224	6256	6320			
6352	6416	6448	6480	6512	6576	6608	6672	6704	6736	6800	6832	6864			
6896	6992	7088	7152	7184	7280	7312	7376	7472	7504	7568	7600	7632			
7664	7760	7792	7856	7920	7952	7984	8048	8112	8144	8176	8208	8240			
8272	8336	8368	8400	8432	8496	8528	8560	8592	8624	8688	8720	8816			
8912	9008	9072	9104	9200	9232	9264	9296	9392	9456	9488	9552	9584			
9616	9648	9680	9712	9776	9872	9968	10032	10064	10128	10160	10192				
10224	10256	10288	10320	10384	10416	10448	10480	10512	10544	10640					
10672	10736	10832	10896	10928	10992	11024	11120	11152	11216	11312					
11344	11376	11408	11440	11472	11536	11600	11632	11696	11728	11792					
11888	11920	11952	11984	12048	12080	12112	12144	12176	12240	12272					
12368	12432	12464	12560	12592	12656	12752	12784	12816	12848	12880					
12912	12944	12976	13040	13072	13136	13168	13232	13264	13328	13392					
13424	13488	13520	13552	13616	13712	13808	13840	13872	13904	14000					
14032	14064	14096	14192	14224	14288	14352	14384	14448	14480	14512					
14544	14576	14608	14640	14672	14736	14768	14832	14864	14896	14960					
14992	15024	15056	15088	15152	15312	15344	15440	15472	15536	15568					
15600	15632	15664	15696	15728	15792	15824	15888	15920	15952	15984					
16016	16112	16208	16272	16304	16400	16432	16496	16592	16656	16720					
16752	16784	16880	16912	16944	16976	17008	17072	17136	17168	17232					
17264	17360	17392	17424	17456	17520	17552	17584	17648	17712	17744					
17808	17840	17872	17936	17968	18032	18064	18096	18128	18192	18224					
18288	18320	18352	18384	18416	18512	18544	18576	18608	18672	18704					
18736	18768	18800	18832	18896	18992	19088	19120	19152	19184	19216					
19312	19344	19376	19408	19472	19568	19600	19632	19664	19696	19760					
19792	19824	19856	19888	19952	20048	20112	20144	20176							

TABLEAU 3

Longueurs  $N/16 \neq 0 \pmod{2}$  pour lesquelles les séquences W.G. n'existent pas $N \leq 20000$ 

226 longueurs existantes

399 longueurs non existantes

TABLEAU 4

Séquences de base de longueur  $N/4 \not\equiv 0 \pmod{2}$   
 $N < 150$

$$N = 12 = 2^2 \times 3 \quad 2 \Phi_{(12)} = 8$$

000100111010

$$N = 20 = 2^2 \times 5 \quad 2 \Phi_{(20)} = 16$$

00101111001101000010

$$N = 28 = 2^2 \times 7 \quad 2 \Phi_{(28)} = 24$$

0110101111100010010100000110

$$N = 36 = 2^2 \times 3^2 \quad 2 \Phi_{(36)} = 24$$

000100001101000100111011110010111010

$$N = 52 = 2^2 \times 13 \quad 2 \Phi_{(52)} = 48$$

0111000111111010100100100010001110000001010110110110

$$N = 60 = 2^2 \times 3 \times 5 \quad 2 \Phi_{(60)} = 32$$

011100010111110010100001001000100011101000001101011110110110

$$N = 76 = 2^2 \times 19 \quad 2 \Phi_{(76)} = 72$$

01110001101011110010111111100100100010001110010101000011010000000111111110

$$N = 84 = 2^2 \times 3 \times 7 \quad 2 \Phi_{(84)} = 48$$

01111101000011010111011100011010000100000001111001001000111001111010110

$$N = 100 = 2^2 \times 5^2 \quad 2 \Phi_{(100)} = 20$$

0111001010110101100011110100100111110111110010001000110101000100111000010110100000110110

$$N = 108 = 2^2 \times 3^3 \quad 2 \Phi_{(108)} = 72$$

0011110001111101110110000001010110001000101001001011001100001100000100010011111101010011101110101010010

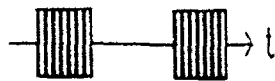
$$N = 124 = 2^2 \times 31 \quad 2 \Phi_{(124)} = 120$$

01100011000010010101011106101000011010000000111010011011000100111001110110101000011010111100101111111000101100100110

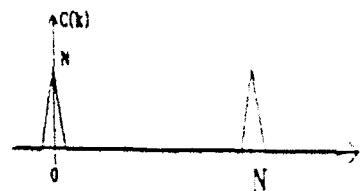
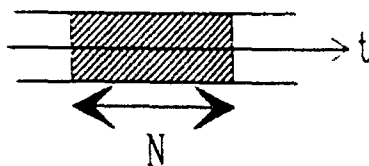
$$N = 148 = 2^2 \times 37 \quad 2 \Phi_{(148)} = 144$$

0110010100010000111110000001001000011001011100010101101011010001000001100010011010111011110000011111011011110011010001110101001010  
1011101111100110

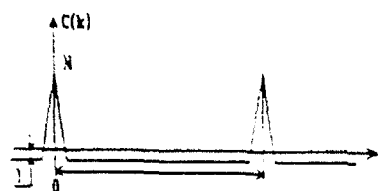
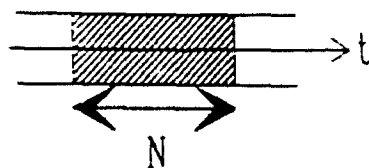
# Sequences de Golay



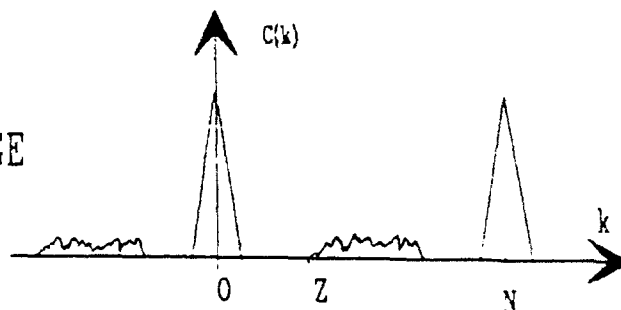
# Sequences multiphase\_Franck\_Sivaswamy\_Goutelard



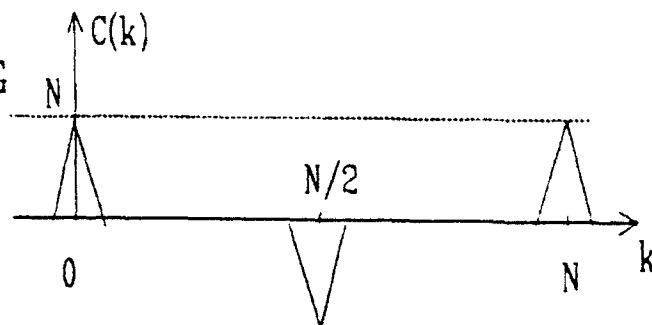
# Sequence binaire PN\_Huffman



# CAZAC - R POLGE



# SEQUENCES W G



**FIGURE 1 :** Séquences parfaites et séquences presque parfaites.



SEQUENCE Q= 1 PERMUTEE/Q'= 1

```
01101011111000
10010100000110
```

SEQUENCE Q= 3 PERMUTEE/Q'= 19

```
01110010000010
10001101111001
```

SEQUENCE Q= 5 PERMUTEE/Q'= 17

```
01100101001111
10011010100000
```

SEQUENCE Q= 9 PERMUTEE/Q'= 25

```
01010011111001
10101000000110
```

SEQUENCE Q= 11 PERMUTEE/Q'= 23

```
00001010001101
11100101110010
```

SEQUENCE Q= 13 PERMUTEE/Q'= 13

```
00100101000001
10011010111110
```

SEQUENCE Q= 15 PERMUTEE/Q'= 15

```
00111110101100
11000001010010
```

SEQUENCE Q= 17 PERMUTEE/Q'= 5

```
00100111010011
11011000101000
```

SEQUENCE Q= 19 PERMUTEE/Q'= 3

```
00110000001010
11001111100101
```

SEQUENCE Q= 23 PERMUTEE/Q'= 11

```
00000010101100
11111001010011
```

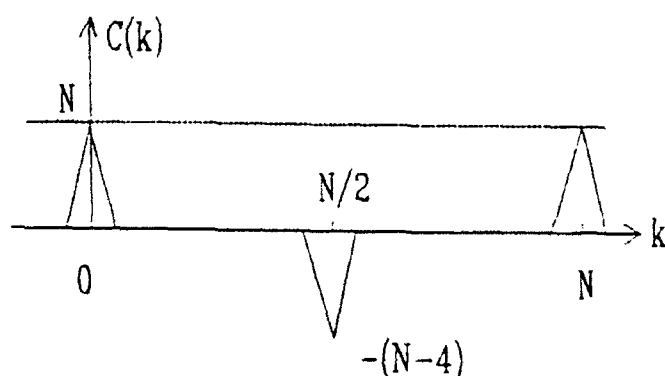
SEQUENCE Q= 25 PERMUTEE/Q'= 9

```
01001111011000
10100000100111
```

SEQUENCE Q= 27 PERMUTEE/Q'= 27

```
00110000010100
10001111101011
```

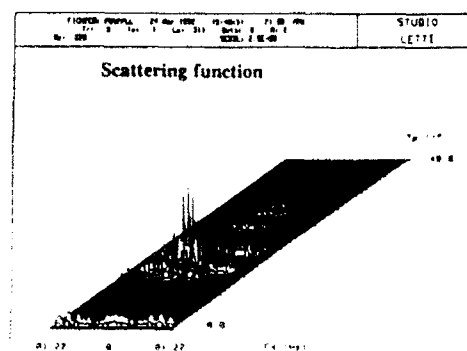
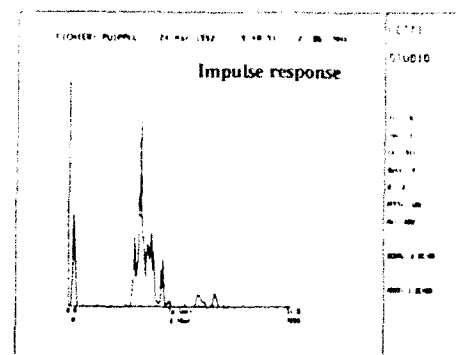
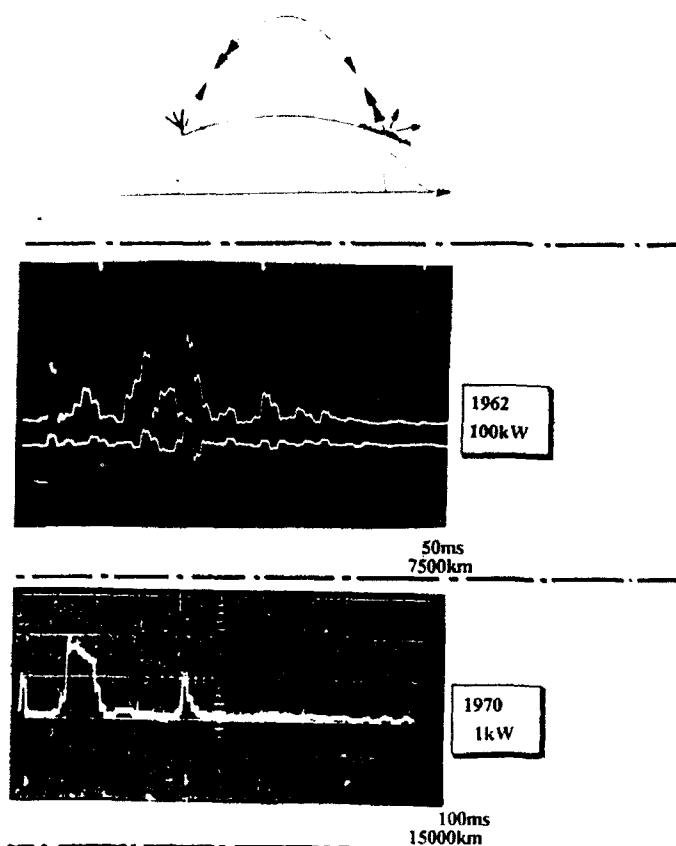
## a) Séquences



## b) Fonction de corrélation

## FIGURE 2

Séquence de base et q permutées  
pour  $N = 28$



1992 : 1W

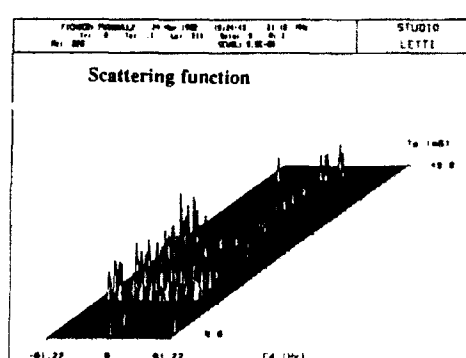
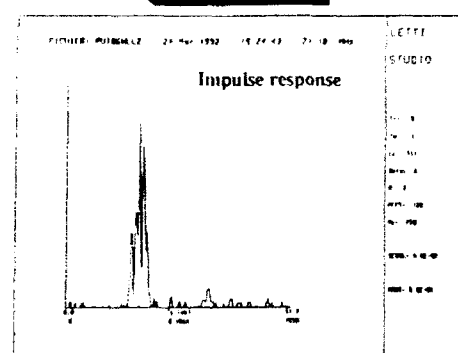
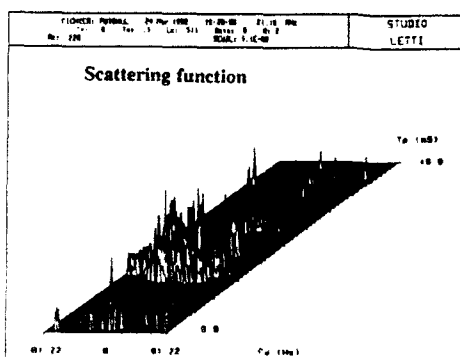
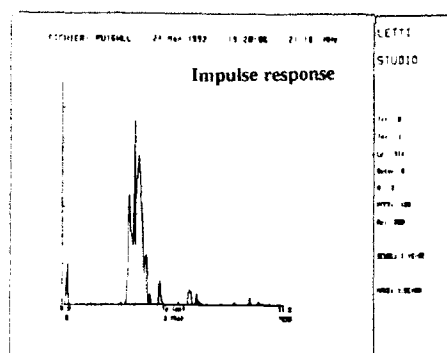


FIGURE 3 : Exemple d'application du codage à la technique "backscatter sounding".  
 Représentation temporelle et fonction de diffusion.

## DISCUSSION

### D. HAINES

Does your method of building up long codes allow the creation of both periodic (repeating 100% duty cycle waveforms) and pulsed codes with perfect autocorrelation functions?

### AUTHOR'S REPLY

Nous utilisons la méthode que nous avons appelée méthode d'impulsions longues codées. Nous avons publié cette technique. Elle consiste à répartir de code sur un ensemble d'impulsions d'émissions entrelacées avec des phases de réception. Un choix judicieux permet d'avoir une couverture exacte de la zone d'observation. Vous pouvez utiliser aussi ces codes dans des émissions continues. *We use a method which we call the encoded long pulse method. We have published this technique. It consists of distributing the code over a set of emission pulses interlaced with reception phases. Careful selection enables correct coverage of the zone of observation. You can also use these codes during continuous emission.*

### L. BERTEL

Pouvez-vous préciser les conditions expérimentales (code, largeur de bande correspondante, durée d'émission) correspondant aux ionogrammes de rétrodiffusion ou fonction de diffusion présentés et relatifs à l'année 1992?

*Can you specify the experimental conditions (code, corresponding bandwidth, duration of emission) for the backscatter ionograms or the scatter function ionograms presented for 1992?*

### AUTHOR'S REPLY

Dans les exemples présentés, pour l'année 1992, les codes utilisés avaient des longueurs voisines de 500 et la bande passante utilisée était de 10kHz. L'émission alternée avec des phases de réception était constituée d'impulsions dont la durée moyenne était de l'ordre de 400 microsecondes.

*In the examples presented for 1992, the lengths of the codes used were around 500 and the bandwidth was 10 kHz. The emission alternating with the reception phases was composed of pulses of an average width of 400 microseconds.*

### F. CHRISTOPHE

Quelles sont les propriétés des séquences Wolfmann-Goutelard en dehors de l'axe des fréquences Doppler nulles, pour ce qui concerne la fonction d'ambiguïté?

*What are the properties of the Wolfmann-Goutelard sequences outside the null Doppler frequency axis in so far as concerns the ambiguity function?*

### AUTHOR'S REPLY

Cette question est très pertinente mais je ne peux y apporter une réponse complète car je n'ai pas encore terminé les calculs relatifs à cette question. Cependant, les caractéristiques de la fonction d'ambiguïté sont liées aux coefficients du  $p$  transformé polynôme du polynôme de corrélation. La régularité de ces coefficients laisse augurer son bon comportement en dehors de l'axe des fréquences Doppler nulles.

*This is a very pertinent question, but unfortunately I cannot give a complete answer, as I have not yet completed the calculations relating to this question. However, the characteristics of the ambiguity function are linked to the coefficients of the polynomial transform of the correlation polynome. The regularity of these coefficients leads us to believe that the behavior is acceptable outside the null Doppler frequency axis.*

# REPORT DOCUMENTATION PAGE

<b>1. Recipient's Reference</b>	<b>2. Originator's Reference</b>	<b>3. Further Reference</b>	<b>4. Security Classification of Document</b>														
	AGARD-CP-528	ISBN 92-835-0695-2	UNCLASSIFIED														
<b>5. Originator</b>	Advisory Group for Aerospace Research and Development North Atlantic Treaty Organization 7 Rue Ancelle, 92200 Neuilly sur Seine, France																
<b>6. Title</b>	RADIOLOCATION TECHNIQUES																
<b>7. Presented at</b>	the Electromagnetic Wave Propagation Panel Symposium, held in London, United Kingdom. 1st-5th June 1992.																
<b>8. Author(s)/Editor(s)</b>	Various		<b>9. Date</b> November 1992														
<b>10. Author's/Editor's Address</b>	Various		<b>11. Pages</b> 354														
<b>12. Distribution Statement</b>	This document is distributed in accordance with AGARD policies and regulations, which are outlined on the back covers of all AGARD publications.																
<b>13. Keywords/Descriptors</b>	<table border="0"> <tr> <td>Radio location</td> <td>Geolocation</td> </tr> <tr> <td>Direction finding</td> <td>HFDF</td> </tr> <tr> <td>Angle of arrival</td> <td>Single site location</td> </tr> <tr> <td>Signal interception</td> <td>Electronic warfare</td> </tr> <tr> <td>Radio interference</td> <td>Interferometry</td> </tr> <tr> <td>Multipath</td> <td>Bearing errors</td> </tr> <tr> <td>Ionospheric variability</td> <td></td> </tr> </table>			Radio location	Geolocation	Direction finding	HFDF	Angle of arrival	Single site location	Signal interception	Electronic warfare	Radio interference	Interferometry	Multipath	Bearing errors	Ionospheric variability	
Radio location	Geolocation																
Direction finding	HFDF																
Angle of arrival	Single site location																
Signal interception	Electronic warfare																
Radio interference	Interferometry																
Multipath	Bearing errors																
Ionospheric variability																	
<b>14. Abstract</b>	<p>Modern defence systems must have a comprehensive surveillance capability. A key component of a surveillance system is the ability to detect and to locate radio signals in the entire radio spectrum from extremely low frequencies through millimetre wavelengths. Various techniques are used to determine the angle of arrival of a radio signal and the distance to the source. This often involves complex signal processing techniques.</p> <p>The symposium addressed first propagation aspects with emphasis on ionospheric characteristics and effects. Two sessions were devoted to the description of radio location techniques such as single site location and another two to radiolocation measurements and systems. One session covered signal processing techniques in radio location applications. One session covered classified subjects and will be published in a classified addendum.</p>																

<p>AGARD Conference Proceedings 528 Advisory Group for Aerospace Research and Development, NATO RADIOLOCATION TECHNIQUES Published November 1992 354 pages</p> <p>Modern defence systems must have a comprehensive surveillance capability. A key component of a surveillance system is the ability to detect and to locate radio signals in the entire radio spectrum from extremely low frequencies through millimetre wavelengths. Various techniques are used to determine the angle of arrival of a radio signal and the distance to the source. This often involves complex signal processing techniques.</p> <p>The symposium addressed first propagation aspects with emphasis on ionospheric characteristics and effects. Two P.T.O.</p>	<p>AGARD-CP-528</p> <p>Radio location Direction finding Angle of arrival Signal interception Radio interference Multipath Ionospheric variability Geolocation HFDF Single site location Electronic warfare Interferometry Bearing errors</p>	<p>AGARD Conference Proceedings 528 Advisory Group for Aerospace Research and Development, NATO RADIOLOCATION TECHNIQUES Published November 1992 354 pages</p> <p>Modern defence systems must have a comprehensive surveillance capability. A key component of a surveillance system is the ability to detect and to locate radio signals in the entire radio spectrum from extremely low frequencies through millimetre wavelengths. Various techniques are used to determine the angle of arrival of a radio signal and the distance to the source. This often involves complex signal processing techniques.</p> <p>The symposium addressed first propagation aspects with emphasis on ionospheric characteristics and effects. Two P.T.O.</p>	<p>AGARD-CP-528</p> <p>Radio location Direction finding Angle of arrival Signal interception Radio interference Multipath Ionospheric variability Geolocation HFDF Single site location Electronic warfare Interferometry Bearing errors</p>
<p>AGARD Conference Proceedings 528 Advisory Group for Aerospace Research and Development, NATO RADIOLOCATION TECHNIQUES Published November 1992 354 pages</p> <p>Modern defence systems must have a comprehensive surveillance capability. A key component of a surveillance system is the ability to detect and to locate radio signals in the entire radio spectrum from extremely low frequencies through millimetre wavelengths. Various techniques are used to determine the angle of arrival of a radio signal and the distance to the source. This often involves complex signal processing techniques.</p> <p>The symposium addressed first propagation aspects with emphasis on ionospheric characteristics and effects. Two P.T.O.</p>	<p>AGARD-CP-528</p> <p>Radio location Direction finding Angle of arrival Signal interception Radio interference Multipath Ionospheric variability Geolocation HFDF Single site location Electronic warfare Interferometry Bearing errors</p>	<p>AGARD Conference Proceedings 528 Advisory Group for Aerospace Research and Development, NATO RADIOLOCATION TECHNIQUES Published November 1992 354 pages</p> <p>Modern defence systems must have a comprehensive surveillance capability. A key component of a surveillance system is the ability to detect and to locate radio signals in the entire radio spectrum from extremely low frequencies through millimetre wavelengths. Various techniques are used to determine the angle of arrival of a radio signal and the distance to the source. This often involves complex signal processing techniques.</p> <p>The symposium addressed first propagation aspects with emphasis on ionospheric characteristics and effects. Two P.T.O.</p>	<p>AGARD-CP-528</p> <p>Radio location Direction finding Angle of arrival Signal interception Radio interference Multipath Ionospheric variability Geolocation HFDF Single site location Electronic warfare Interferometry Bearing errors</p>

<p>sessions were devoted to the description of radio location techniques such as single site location and another two to radiolocation measurements and systems. One session covered signal processing techniques in radio location applications. One session covered classified subjects and will be published in a classified addendum.</p> <p>Papers presented at the Electromagnetic Wave Propagation Panel Symposium, held in London, United Kingdom, 1st—5th June 1992.</p> <p>ISBN 92-835-0695-2</p>	<p>sessions were devoted to the description of radio location techniques such as single site location and another two to radiolocation measurements and systems. One session covered signal processing techniques in radio location applications. One session covered classified subjects and will be published in a classified addendum.</p> <p>Papers presented at the Electromagnetic Wave Propagation Panel Symposium, held in London, United Kingdom, 1st—5th June 1992.</p> <p>ISBN 92-835-0695-2</p>
<p>sessions were devoted to the description of radio location techniques such as single site location and another two to radiolocation measurements and systems. One session covered signal processing techniques in radio location applications. One session covered classified subjects and will be published in a classified addendum.</p> <p>Papers presented at the Electromagnetic Wave Propagation Panel Symposium, held in London, United Kingdom, 1st—5th June 1992.</p> <p>ISBN 92-835-0695-2</p>	<p>sessions were devoted to the description of radio location techniques such as single site location and another two to radiolocation measurements and systems. One session covered signal processing techniques in radio location applications. One session covered classified subjects and will be published in a classified addendum.</p> <p>Papers presented at the Electromagnetic Wave Propagation Panel Symposium, held in London, United Kingdom, 1st—5th June 1992.</p> <p>ISBN 92-835-0695-2</p>

AGARD

NATO OTAN

7 RUE ANCELLE · 92200 NEUILLY-SUR-SEINE

FRANCE

Téléphone (1)47.38.57.00 · Télex 610 176

Télécopie (1)47.38.57.99

DIFFUSION DES PUBLICATIONS

AGARD NON CLASSIFIEES

L'AGARD ne détient pas de stocks de ses publications, dans un but de distribution générale à l'adresse ci-dessus. La diffusion initiale des publications de l'AGARD est effectuée auprès des pays membres de cette organisation par l'intermédiaire des Centres Nationaux de Distribution suivants. A l'exception des Etats-Unis, ces centres disposent parfois d'exemplaires additionnels; dans les cas contraire, on peut se procurer ces exemplaires sous forme de microfiches ou de microcopies auprès des Agences de Vente dont la liste suit.

CENTRES DE DIFFUSION NATIONAUX

**ALLEMAGNE**

Fachinformationszentrum,  
Karlsruhe  
D-7514 Eggenstein-Leopoldshafen 2

**BELGIQUE**

Coordonnateur AGARD-VSL  
Etat-Major de la Force Aérienne  
Quartier Reine Elisabeth  
Rue d'Evere, 1140 Bruxelles

**CANADA**

Directeur du Service des Renseignements Scientifiques  
Ministère de la Défense Nationale  
Ottawa, Ontario K1A 0K2

**DANEMARK**

Danish Defence Research Board  
Ved Idraetsparken 4  
2100 Copenhagen Ø

**ESPAGNE**

INTA (AGARD Publications)  
Pintor Rosales 34  
28008 Madrid

**ETATS-UNIS**

National Aeronautics and Space Administration  
Langley Research Center  
M/S 180  
Hampton, Virginia 23665

**FRANCE**

O.N.E.R.A. (Direction)  
29, Avenue de la Division Leclerc  
92322 Châtillon Cedex

**GRECE**

Hellenic Air Force  
Air War College  
Scientific and Technical Library  
Dekelia Air Force Base  
Dekelia, Athens TGA 1010

**ISLANDE**

Director of Aviation  
c/o Flugrad  
Reykjavik

**ITALIE**

Aeronautica Militare  
Ufficio del Delegato Nazionale all'AGARD  
Aeroporto Pratica di Mare  
00040 Pomezia (Roma)

**LUXEMBOURG**

Voir Belgique

**NORVEGE**

Norwegian Defence Research Establishment  
Attn: Biblioteket  
P.O. Box 25  
N-2007 Kjeller

**PAYS-BAS**

Netherlands Delegation to AGARD  
National Aerospace Laboratory NLR  
Kluyverweg 1  
2629 HS Delft

**PORTUGAL**

Portuguese National Coordinator to AGARD  
Gabinete de Estudos e Programas  
CLAFIA  
Base de Alfragide  
Alfragide  
2700 Amadora

**ROYAUME UNI**

Defence Research Information Centre  
Kentigern House  
65 Brown Street  
Glasgow G2 8EX

**TURQUIE**

Milli Savunma Başkanlığı (MSB)  
ARGE Daire Başkanlığı (ARGE)  
Ankara

LE CENTRE NATIONAL DE DISTRIBUTION DES ETATS-UNIS (NASA) NE DETIENT PAS DE STOCKS  
DES PUBLICATIONS AGARD ET LES DEMANDES D'EXEMPLAIRES DOIVENT ETRE ADRESSEES DIRECTEMENT  
AU SERVICE NATIONAL TECHNIQUE DE L'INFORMATION (NTIS) DONT L'ADRESSE SUIT.

AGENCES DE VENTE

National Technical Information Service  
(NTIS)  
5285 Port Royal Road  
Springfield, Virginia 22161  
Etats-Unis

ESA/Information Retrieval Service  
European Space Agency  
10, rue Mario Nikis  
75015 Paris  
France

The British Library  
Document Supply Division  
Boston Spa, Wetherby  
West Yorkshire LS23 7BQ  
Royaume Uni

Les demandes de microfiches ou de photocopies de documents AGARD (y compris les demandes faites auprès du NTIS) doivent comporter la dénomination AGARD, ainsi que le numéro de série de l'AGARD (par exemple AGARD-AG-315). Des informations analogues, telles que le titre et la date de publication sont souhaitables. Veuillez noter qu'il y a lieu de spécifier AGARD-R-*nnn* et AGARD-AR-*nnn* lors de la commande de rapports AGARD et des rapports consultatifs AGARD respectivement. Des références bibliographiques complètes ainsi que des résumés des publications AGARD figurent dans les journaux suivants:

Scientific and Technical Aerospace Reports (STAR)  
publié par la NASA Scientific and Technical  
Information Division  
NASA Headquarters (NTT)  
Washington D.C. 20546  
Etats-Unis

Government Reports Announcements and Index (GRA&I)  
publié par le National Technical Information Service  
Springfield  
Virginia 22161  
Etats-Unis  
(accessible également en mode interactif dans la base de  
données bibliographiques en ligne du NTIS, et sur CD-ROM)



Imprimé par Specialised Printing Services Limited  
40 Chigwell Lane, Loughton, Essex IG10 3TZ